DOCUMENT RESUME

ED 056 722 LI 003 231

AUTHOR Raitt, D. I.
TITLE Space Documentation Services: Operations Handbook.
INSTITUTION European Space Research Organization, Neuilly-Sur-Seine (France).
REPORT NO ESRO-SP-63
PUB DATE May 71
NOTE 44p.; (0 References)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Aerospace Technology; Documentation; Electronic Data Processing; Foreign Countries; Information Needs; Information Processing; *Information Retrieval; *Information Services; Man Machine Systems; Manuals; *On Line Systems; *Search Strategies
IDENTIFIERS SDS; *Space Documentation Service

ABSTRACT

A description of the system used by Space Documentation Service (SDS) to disseminate combined and accumulated knowledge, as widely as possible, throughout Europe is given. The RECON network, with the full support of NASA, has gradually been extended so that centers in Member States, may, by installing their own terminals, have direct access to the data base. The data base itself has been broadened to include not only the NASA file, but also that of Metals Abstracts, Engineering Index and the U. S. Government Research and Development Reports. Other files are under consideration. The feasibility of developing new features which would be of value in the searching of the data-base is being studied and a continuous monitoring of the system, services and procedures is taking place to ensure efficiency and effectiveness. (Author)

EDO 56722

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

R.N.
ESRO SP-63

SPACE DOCUMENTATION SERVICE : OPERATIONS HANDBOOK

by D.I. Raitt

May 1971

ORGANISATION EUROPÉENNE DE RECHERCHES SPATIALES
EUROPEAN SPACE RESEARCH ORGANISATION,
114, Avenue Charles-de-Gaulle, 92 - NEUILLY-SUR-SEINE (France)

LI 003 231

TABLE OF CONTENTS

INTRODUCTION 5

Information needs; automated systems; advantages of on-line systems; summary.

EUROPE'S RETRIEVAL SYSTEM 7

The Space Documentation Service; how information is recorded; how information is announced; retrieving the information; RECON.

SEARCH ROUTINE 11

The search; operations; output presentation.

QUESTION FORMULATION 15

User interaction; question; purpose; references; keywords; files; limitations; request forms.

SEARCH STRATEGIES 19

The coordination problem; search logic; vocabularies; generic relationships; conceptual relationships; concept grouping; omission of implicit categories; term selection based on display; search modification; aspect searching; different approaches; reminders.

RECON HARDWARE AND SOFTWARE 29

The RECON network; terminal equipment; computer equipment; storage media; the RECON system; update procedure; SDI procedure.

DEVELOPMENT OF NEW FEATURES FOR RECON 33

Improvements to Boolean formulations; free language formulation; term extraction and frequency of occurrence; document ranking; keeping references; file of related terms; output presentation.

RECON COMMANDS 37

CONCLUSION 45

INTRODUCTION

Information needs

Information has no intrinsic value until it reaches the point at which it can be acted upon. In the complex environment of today it is generally accepted that people asking for information, be they scientists, engineers, technologists, managers or just laymen, need it quickly, accurately and in an organised manner if time and money are not to be wasted. An organisation cannot produce more than a small part of the information it needs, and thus there must be an effective mechanism for coupling it to a much larger continually expanding store of information that is available to all. There is a definite need for a rapid, efficient and correct information transfer but normal manual information retrieval systems have often a high time lapse and a low hit response, although both can be and are being overcome by the use of automated systems.

Automated systems

One of the main reasons for using an automated or computer system is the ability of the machine to do logical detailed searches at a far greater speed than its laborious manual counterpart. Many of these computerized retrieval systems, however, suffer an inherent disadvantage of providing no intermediate results for user evaluation and subsequent modification of search strategy, and though the system is automated, access is slow and time is lost due to off-line and batch-mode processing.

(a) *Batch processing*

Even though a system may possess a computer which can do its work at a fantastic speed, the results are still probably not being received quickly enough by those who want them.

This is because the data and requests take a long time to reach the computer due to the necessity of punching cards or tape. Then since the computer is prohibitively costly and time-consuming, data to be put into the system is often queued and batched and all subsequent operations are done on that batch as a whole i.e. the batch may be scanned and only when that operation is complete, the batch might be merged or sorted and so on. In ordinary operations, hours sometimes days between batches may elapse. If the information is in a batched mode on magnetic tape, the file will have to be searched sequentially to find the relevant items. Finally, there is the delay in getting the results back to the enquirer, since he is usually some distance from the computer and the information has to first pass through an intermediary.

(b) *On-line systems*

To overcome these drawbacks systems described as « on-line » and « real-time » are now installed in many organisations, and they allow a user to converse directly with the computer in his search for information. The user, remote from the central processor, can by means of a keyboard and display screen, submit a request for information to the computer over telecommunication lines. The central processor will search its files for the relevant information, perform any necessary computations and transmit the results of the search back to the remote user via the display screen and teletype. This facility means that libraries and information centres by utilizing this new technology can have easier and quicker access to a store of information, than they could otherwise have had with a purely manual system.

Advantages of on-line systems

There are many situations requiring on-line control e.g. banking, airline reservations, stock control, air traffic control, where it is essential to have completely up-to-date information quickly so that decisions can be made.

In an information context the use of an integrated remote access retrieval system which is coupled directly to a computer will provide the library with immediate access to all the information stored in the computer. The concept of immediate access is not new since the catalogues employed in most libraries are based on this principle. Thus to improve library services by the utilization of an automated system, access to the file must still be immediate and the development of an on-line retrieval system is one method of improvement.

One of the main advantages of such a system is that it is specifically designed to provide a man-machine interface. Programmes can be written in such a way that an inexperienced operator can communicate with the computer using natural language, will have his actions checked by the computer and will be informed of his errors by the machine. By the elimination of intermediaries and placing the user in direct contact with the computer, results are made available in seconds instead of days and search interests can be redefined based on an examination of the intermediate result.

Another advantage is that the number of references on a given topic no matter how specialised may be obtained immediately and these can be scanned on the display screen to enable their relevance to be assessed, and to provide other likely avenues of exploration. There is the capability of simply browsing, with any useful references being kept and later printed out. The speed of response, compared to off-line or batched systems means that a user can follow his chain of thought without interruptions and at his own speed.

One further advantage is that several consoles can be coupled to the computer enabling several users to access and search the files at the same time. It is clear that such a system will avoid the time-consuming manual searching of catalogues and abstract journals and the physical checking of documents and also the delays to which off-line systems are subjected.

Summary

The effectiveness of a remote-access system then lies in the degree to which it can accomplish the following:

- Minimize response time i.e. the time between the receipt of a request by the library and the receipt of the information by the enquirer.
- Allow evaluation of intermediate results and subsequent modification of search strategy.
- Provide a flexible programme with command functions and operational modes capable of being understood by an inexperienced operator and modified according to his needs.
- Allow the control of the complete search operation to lie with the user.
- Eliminate the need for an intermediary and so allow a real-time man-machine interface.

If the system can accomplish these basic requirements then a user has literally at his finger tips the inspiring potential of coupling the computer's fantastic memory and calculating speed to his own intellect, speed and immediate requirements for information.

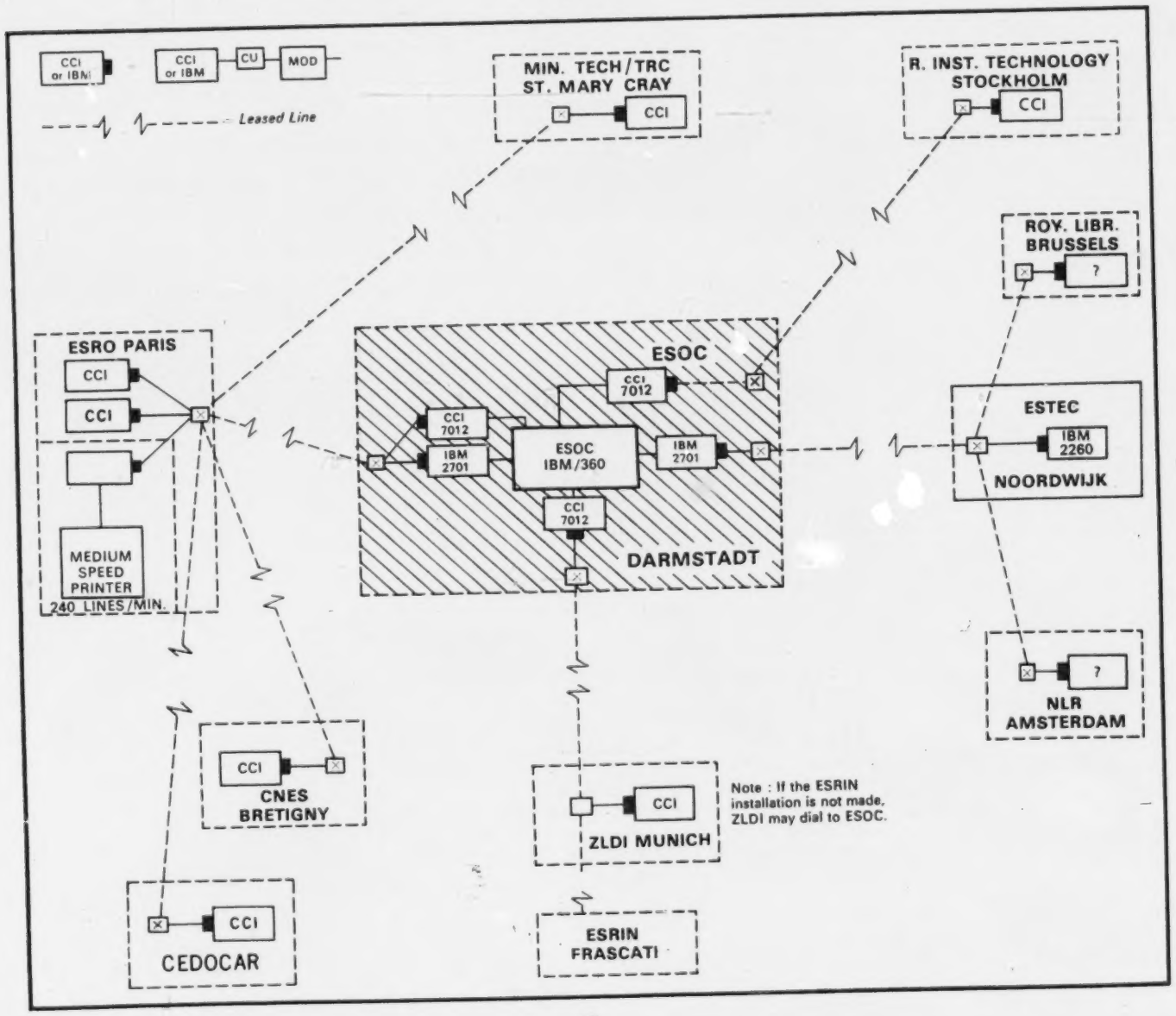
EUROPE'S RETRIEVAL SYSTEM

The Space Documentation Service

The Space Documentation Service (SDS), set up in 1965 by the European Space Research Organisation (ESRO) and the European Space Vehicle Launcher Development Organisation (ELDO), is part of a comprehensive information programme to make available in Europe the results of research and development effort in the field of science and technology.

Until recently SDS, situated in Paris, employed the slow off-line process whereby search requests were sent to a contractor's computer in America for processing. A search consisted of running through the whole file seeking matches between the search argument and the indexing terms. Questions were batched to save computer time (and thus costs) but the search time was

Fig. 1



still about 12 hours. This, together with the postal time taken to send data to and forth across the Atlantic, meant quite considerable delays.

SDS has successfully overcome the problems and delays outlined above by the installation of an on-line, remote access retrieval system (NASA/RECON) which is the first of its kind in Europe.

The basis of the system is a generous exchange agreement by which the National Aeronautics and Space Administration (NASA) makes available to SDS information published in STAR (Scientific and Technical Aerospace Reports) and IAA (International Aerospace Abstracts) since 1962, in magnetic tape form, and also makes available microfiches of the documents for dissemination. In return, SDS indexes, abstracts and supplies European publications to NASA for inclusion in the system. The prompt, comprehensive and effective dissemination of this knowledge in Europe is the aim of the ESRO/ELDO SDS, and the installation of the new high speed information retrieval system has improved European capabilities for the rapid searching of literature, since the system has been extended to provide direct access to the computer for national centres in Member States (Fig. 1).

How information is recorded

Users of SDS, i.e. research institutes, universities, firms, in ESRO Member States supply copies of their reports to SDS. These are promptly checked to avoid duplication, examined for relevance, and then catalogued descriptively. The documents are then indexed in depth, using keywords from the NASA Thesaurus, and an abstract is prepared if one does not already exist. The bibliographic information, together with the document and microfiche copy, is sent to NASA who assign a unique accession number to each document so that it can always be readily traced. The information is then stored at the NASA Facility, with other non-ESRO material.

How information is announced

The descriptions of the documents together with their accession numbers appear in one of two semi-monthly abstracting and indexing journals. (Fig. 2). STAR covers the world's unpublished (report) literature in the broad aerospace field and IAA covers the world's published (periodicals, conferences) literature in the same field. Each journal appears in alternate weeks and provides a comprehensive access to current literature on aerospace science and technology.

N70-16144# Deutsche Forschungsanstalt fur Luft- und Raumfahrt, Brunswick (West Germany). Inst. fuer Strahlantriebe.
CROSSED FIELD ACCELERATION OF A MULTI-GAS PLASMA [KREUZFELDBESCHLEUNIGUNG DES MEHRGAS-PLASMAS]
Gisela Stoffers Jan. 1969 26 p refs In GERMAN; ENGLISH summary.
(DLR-FB-69-45) Avail: CFSTI.

The plasma is treated as a system of several gases flowing in the same space with different velocities and not in thermodynamic equilibrium with each other. Friction forces occurring between the gases are assumed to be continuously distributed in space. For the evaluation of the friction process between two of the gases, their average velocity is defined. The friction force on one of the two gases is proportional to the difference between its own velocity and the average velocity. The technical work done per time unit on the gas, is proportional to the product of the above velocity difference and the average velocity. The frictional work per time unit is proportional to the square product of that difference.
Author (ESRO)

Fig. 2

Retrieving the information

All references appearing in STAR and IAA are stored in a direct access data cell which now contains nearly 500,000 references and is being added to at the rate of 6-7,000 items per month. Every month, a tape containing the new references is sent by NASA to the European Space Operations Centre (ESOC) in Darmstadt, Germany, where it is loaded into the computer store to join all the other references since 1962. Thus, by accessing this store by the NASA keywords used in indexing, the unique accession numbers, authors or corporate sources it is possible to provide retrospective bibliographical searches or current awareness searches. By making use of RECON, SDS is able to retrieve references and offer these services with a high degree of relevance being assured and with a minimum delay.

RECON (Remote Console)

RECON is an interactive information retrieval system, developed by Lockheed Aircraft Corporation for NASA, which allows the user to formulate simple or complex search requests via a video keyboard display terminal (Fig. 3) which is coupled to the ESRO computer (IBM 360 65) in Darmstadt (Germany). To use RECON the analyst enters desired commands such as « BEGIN »,

Fig. 3

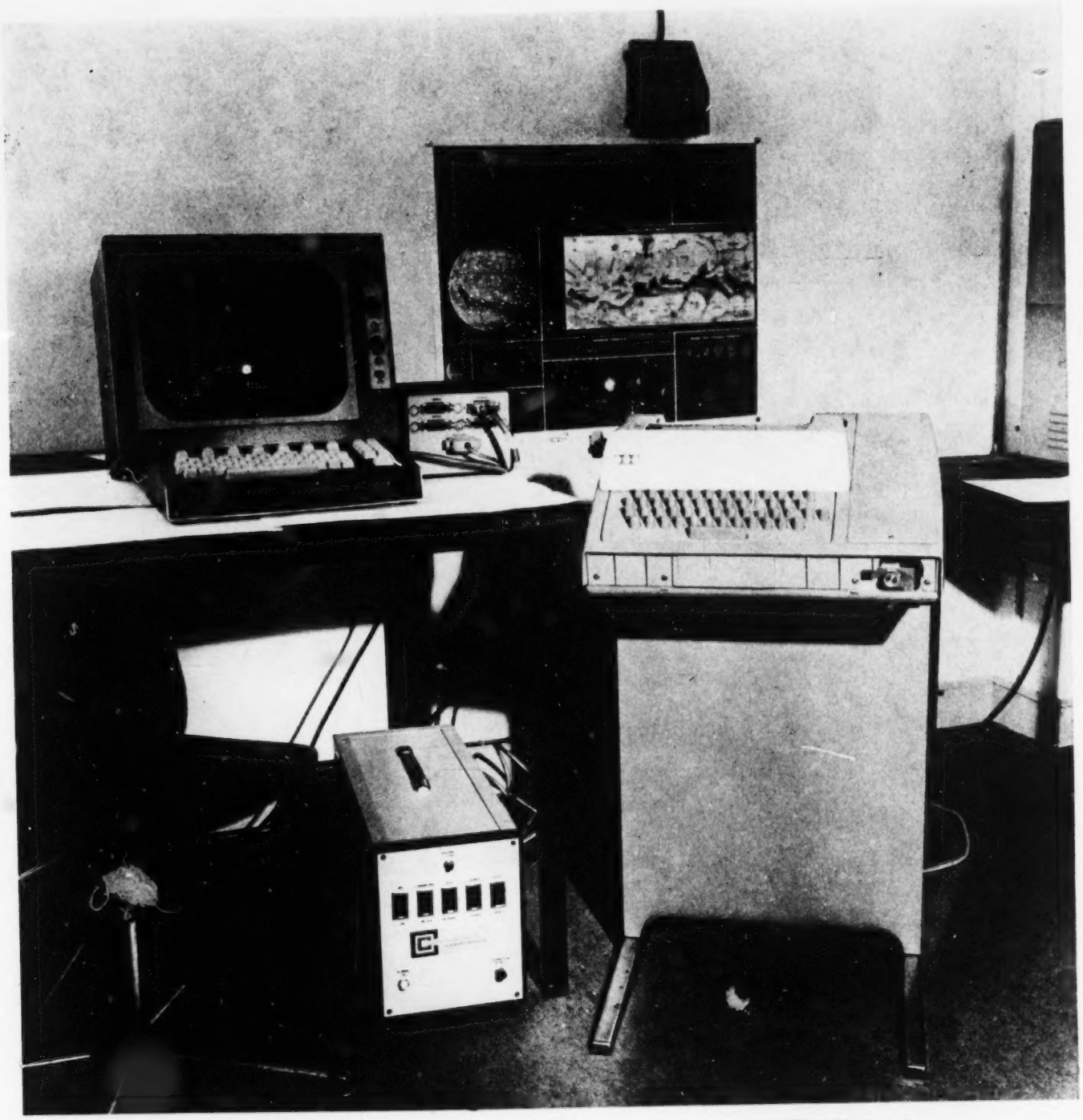


Fig. 4

« SELECT », « EXPAND », « COMBINE » or « DISPLAY » by depressing keys on the keyboard (Fig. 4). The computer responds instantaneously by displaying various data on the console screen and the teletype which is connected to it.

There are four basic phases in the search expression:

- the selection of index terms
- the combination of these terms
- the displaying of results
- the modification of the search strategy.

These four phases form an interactive loop.

In the first step, the analyst identifies and selects index terms using the controlled vocabulary of the NASA Thesaurus which relate to the search topic. The second step allows the user to build up his own specialized subject by forming logical equations based on Boolean Algebra. At each step the analyst is informed of the number of documents which fulfill his specified search criteria. The third step, display of output, allows the analyst to review references for relevancy and to select additional terms based on index terms found in displayed references to better express his search requirement (fourth step). This latter step is a feedback step which closes the loop (Fig. 5).

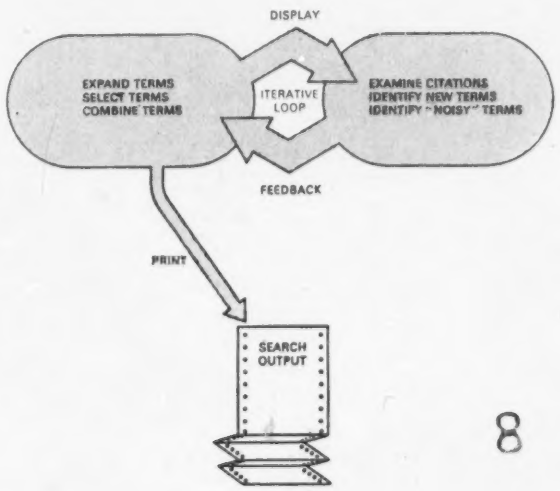


Fig. 5

8

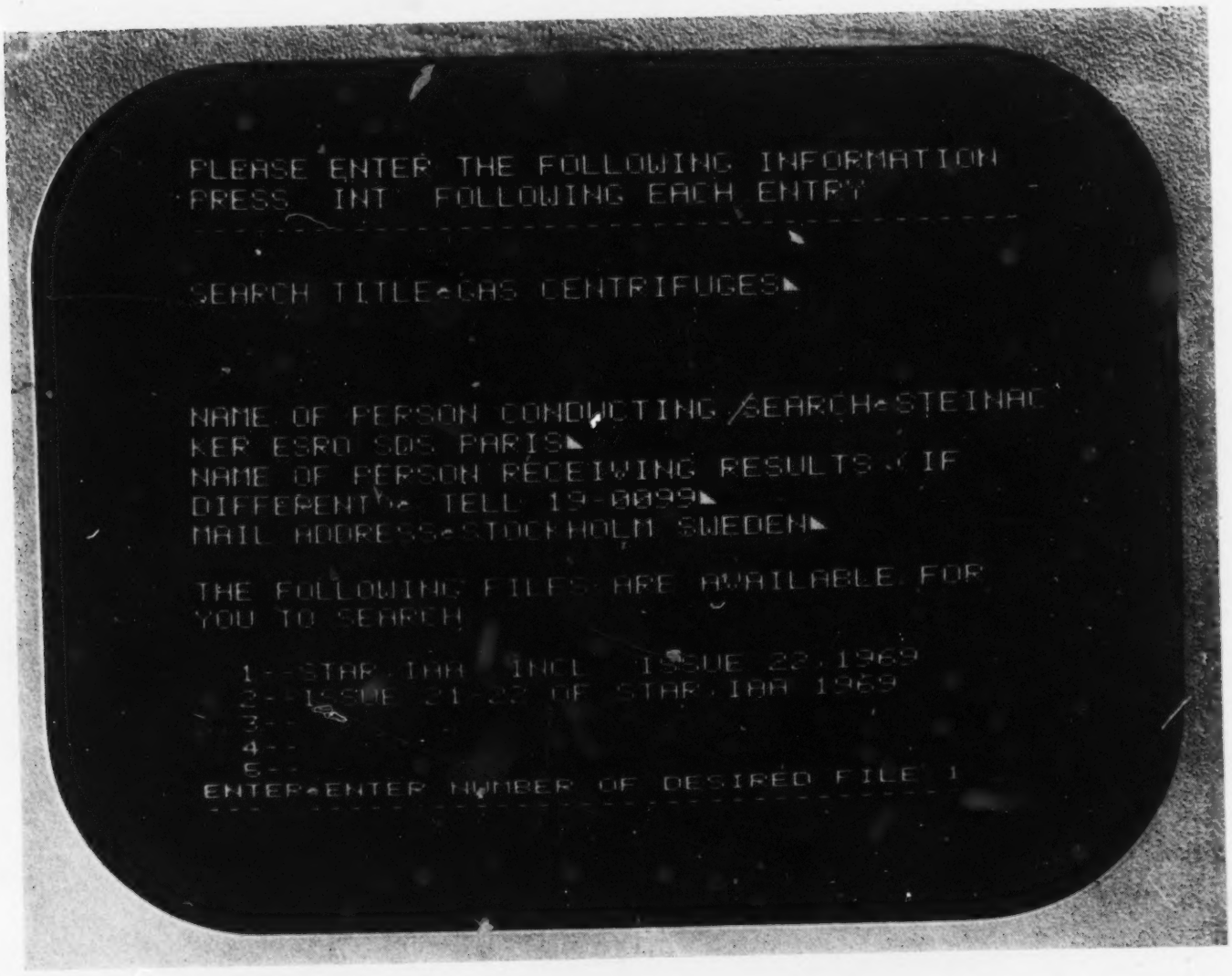
SEARCH ROUTINE

The search

Let us take, by way of example, a search for information on Gas Centrifuges, and assume that the user is conversant with the basic principles outlined above. His first action is to press the « BEGIN SEARCH » key, the result of which is illustrated in the first part of the accompanying photographs showing four separate displays on the screen at different stages of the search. This first illustration (Fig. 6) shows the computer asking a number of routine questions such as title of search, name of searcher, etc. The information given in reply to the questions serves to identify the subsequent output.

Then the actual search strategy takes place, as described below. At the end of the search process the command « END SEARCH » has to be given. This will also allow the analyst to express some comments on the search and on the system performance in general. The time elapsed between the two commands « BEGIN » and « END » is the search time, or console time. The computer (CPU) time is a mere fraction of this.

Fig. 6



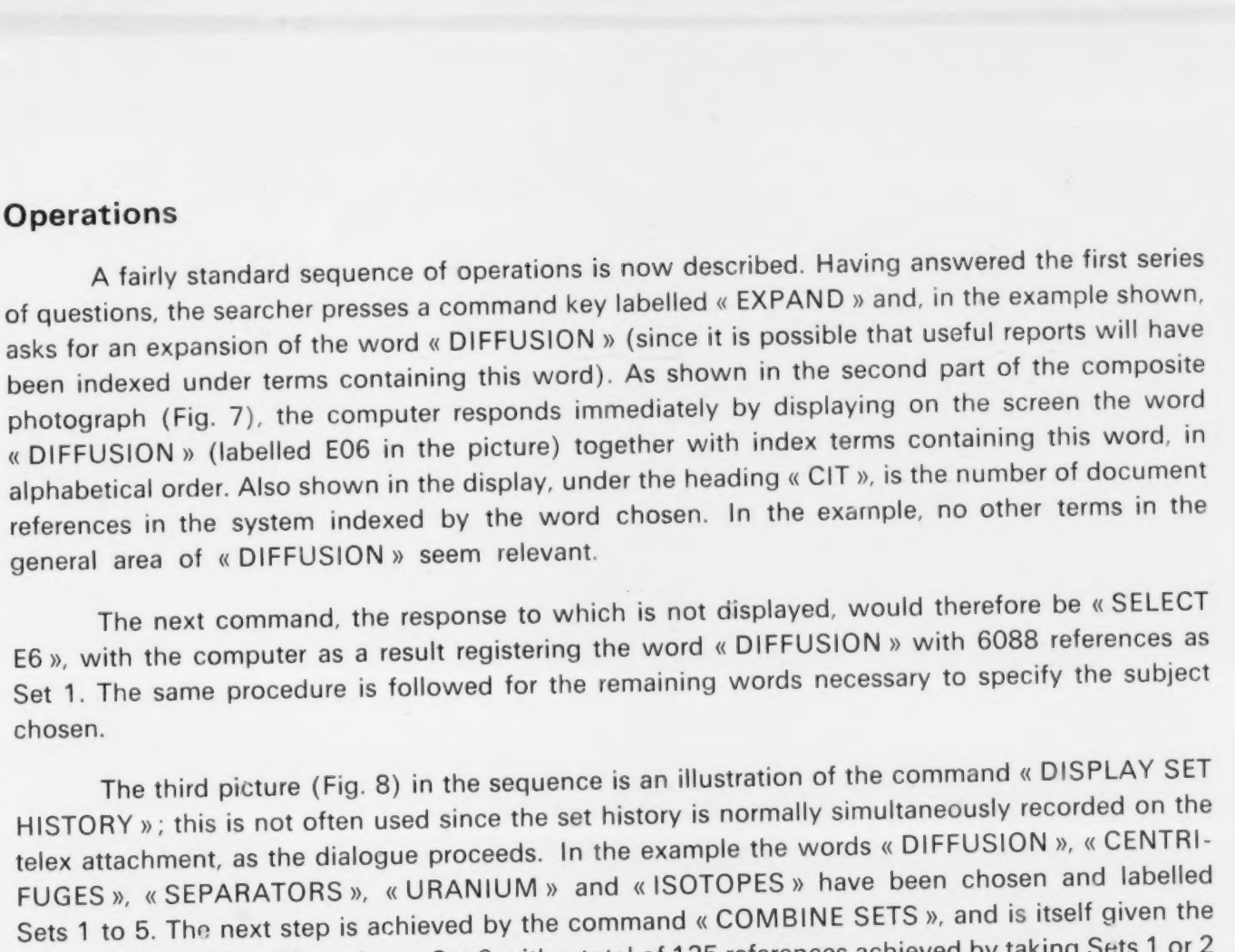
Operations

A fairly standard sequence of operations is now described. Having answered the first series of questions, the searcher presses a command key labelled « EXPAND » and, in the example shown, asks for an expansion of the word « DIFFUSION » (since it is possible that useful reports will have been indexed under terms containing this word). As shown in the second part of the composite photograph (Fig. 7), the computer responds immediately by displaying on the screen the word « DIFFUSION » (labelled E06 in the picture) together with index terms containing this word, in alphabetical order. Also shown in the display, under the heading « CIT », is the number of document references in the system indexed by the word chosen. In the example, no other terms in the general area of « DIFFUSION » seem relevant.

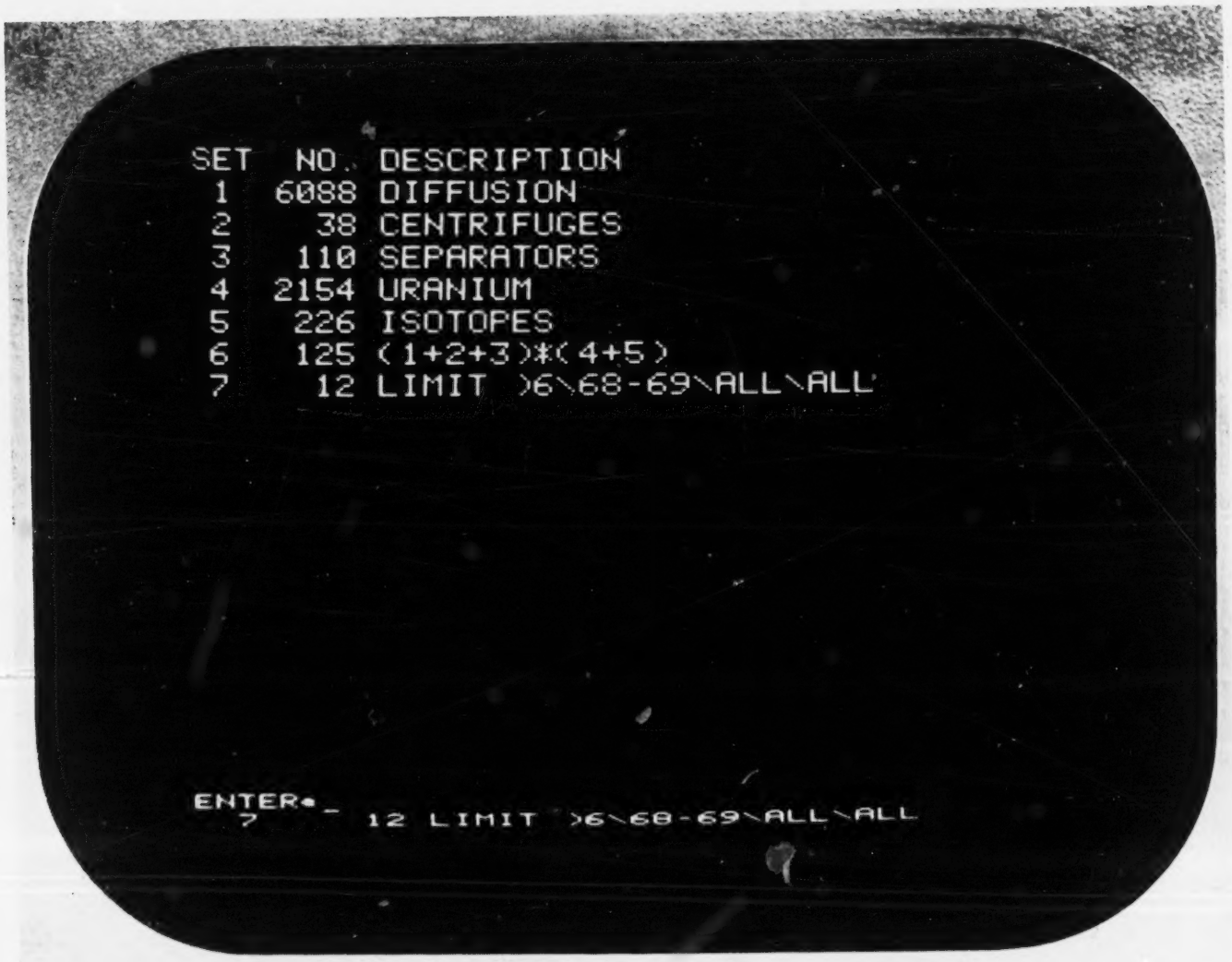
The next command, the response to which is not displayed, would therefore be « SELECT E6 », with the computer as a result registering the word « DIFFUSION » with 6088 references as Set 1. The same procedure is followed for the remaining words necessary to specify the subject chosen.

The third picture (Fig. 8) in the sequence is an illustration of the command « DISPLAY SET HISTORY »; this is not often used since the set history is normally simultaneously recorded on the telex attachment, as the dialogue proceeds. In the example the words « DIFFUSION », « CENTRIFUGES », « SEPARATORS », « URANIUM » and « ISOTOPES » have been chosen and labelled Sets 1 to 5. The next step is achieved by the command « COMBINE SETS », and is itself given the reference Set 6. The screen shows Set 6 with a total of 125 references achieved by taking Sets 1 or 2 or 3 each in combination with Sets 4 or 5. In other words only reports which have been indexed under the terms « DIFFUSION » and « URANIUM » or « DIFFUSION » and « ISOTOPES », etc. will be selected from the file of half a million document references.In Set 7 the effect of the command « LIMIT », in this case to references received in 1968 and 1969, may be seen. The 125 references have now been reduced to 12.

Fig. 7

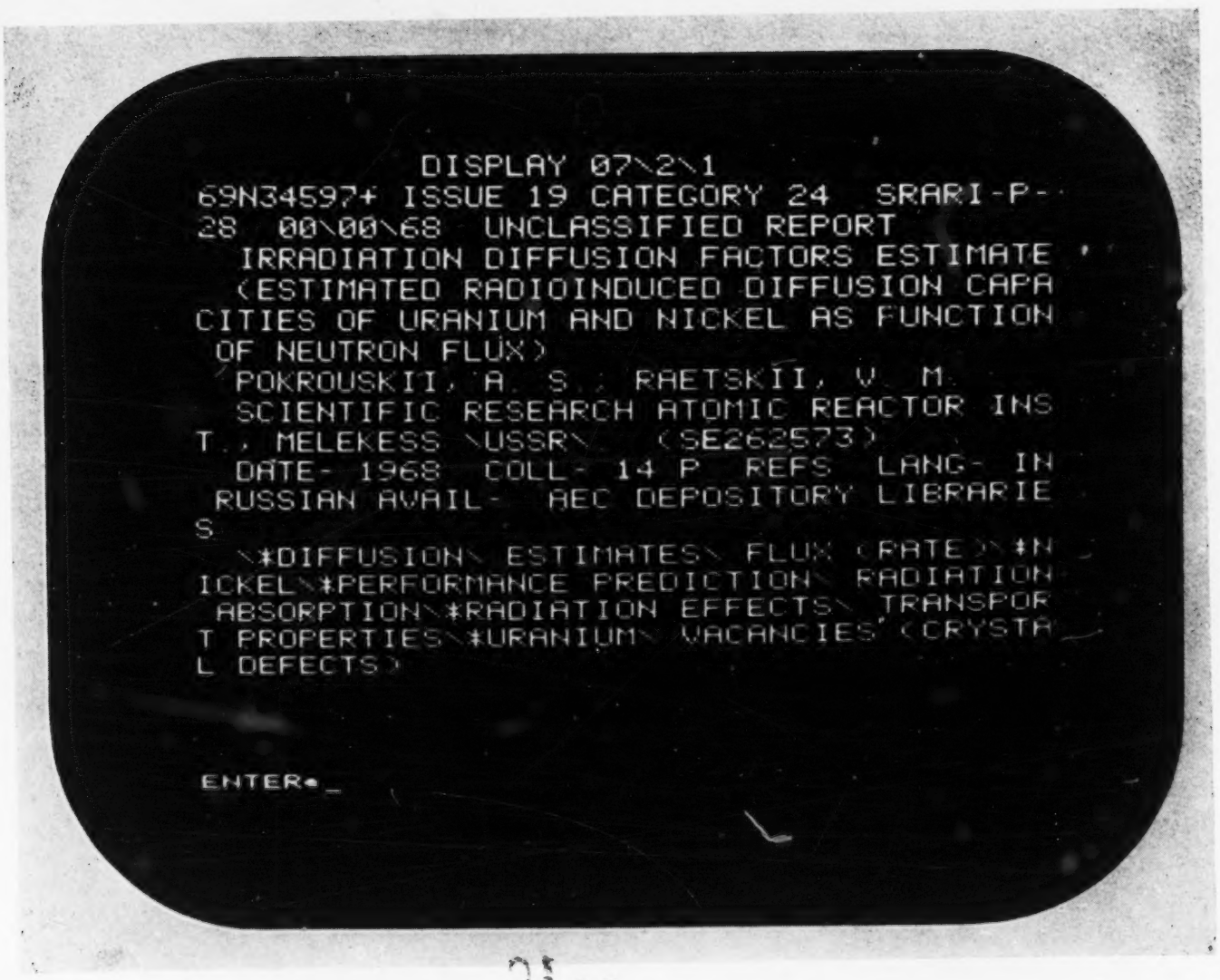


10


SET NO. DESCRIPTION
1 6088 DIFFUSION
2 38 CENTRIFUGES
3 110 SEPARATORS
4 2154 URANIUM
5 226 ISOTOPES
6 125 (1+2+3)*(4+5)
7 12 LIMIT)6\68-69\ALL\ALL
ENTER•_
7 12 LIMIT)6\68-69\ALL\ALL

▲ *Fig. 8*

Fig. 9 ▼


DISPLAY 07\2\1
69N34597+ ISSUE 19 CATEGORY 24 SRARI-P-28 00\00\68 UNCLASSIFIED REPORT
IRRADIATION DIFFUSION FACTORS ESTIMATE (ESTIMATED RADIOINDUCED DIFFUSION CAPACITIES OF URANIUM AND NICKEL AS FUNCTION OF NEUTRON FLUX)
POKROUSKII, A. S.; RAETSKII, V. M.
SCIENTIFIC RESEARCH ATOMIC REACTOR INST.; MELEKESS \USSR\ (SE262573)
DATE- 1968 COLL- 14 P REFS LANG- IN RUSSIAN AVAIL- AEC DEPOSITORY LIBRARIES
*DIFFUSION\ ESTIMATES\ FLUX (RATE)\ *NICKEL*PERFORMANCE PREDICTION\ RADIATION ABSORPTION*RADIATION EFFECTS\ TRANSPORT PROPERTIES*URANIUM\ VACANCIES (CRYSTAL DEFECTS)
ENTER•_

In the fourth photograph (Fig. 9) the use of the « DISPLAY » command is shown. Each of the 12 references selected may be examined on the screen in the same way. As can be seen, the title of the report is followed by, in brackets, an « abstract of an abstract » or notation of content. As the index terms are also displayed on the screen, further refinement of the logical equation in Set 6 may be undertaken. If necessary, each of the records can be inspected before printing on the high-speed printer, so that an editing function may be carried out using the command « KEEP ». The average time for a search is about 30 minutes.

Output presentation

The results of the above operation are printed on a high-speed printer (1,200 lines/minutes) by giving the « PRINT » command, the output being normally exactly the same as that shown in the photograph. It is possible, however, to give the command « TYPE » with the result that the references are printed on the local printer. Other facilities are also available such as, for example, the ability to pass messages, via the screen, between terminals.

The power of the system may also be extended by the addition of new programme modules. In fact, each command instruction within the RECON software is achieved by a separate programme module. Thus it is not too difficult to add new modules as required.

QUESTION FORMULATION

User's interaction

The RECON system is a powerful device, but its performance depends mostly upon the information supplied by the user in the definition of his question. The following sections are intended to aid users in the formulation of questions to be processed by SDS so that they may get the maximum response from the system with the minimum trouble. The more help that can be given before the search, the better the search will be. SDS would find it most helpful if the following information could be given when a computer search is requested.

Question

The construction of the search formula is a translative one, converting the question as posed by the user in his own language into the retrieval language employed by the system. As long as idiosynchrasies of machine-manipulated natural language remain largely unknown, questions will very often be formulated in a manner highly unsuitable for correct translation.

It is better to write the question in a short paragraph, rather than just give odd words to be logically combined. It is also important that the most precise terms possible are used. In particular it must be borne in mind that a search on the term « alloys » will not automatically produce all the different types of alloys such as aluminium or titanium alloys. These must be searched on as separate keywords. In view of the large number of citations to which terms are indexed, it is essential to specify exactly which alloys, materials, properties or whatever are required. If all types of alloys *are* required, then this should be indicated, in which case it is likely that the search would have to be limited by date to avoid running into shortage of computer storage space.

Below are some typical questions as they were submitted by users. A revised statement is given on the right:

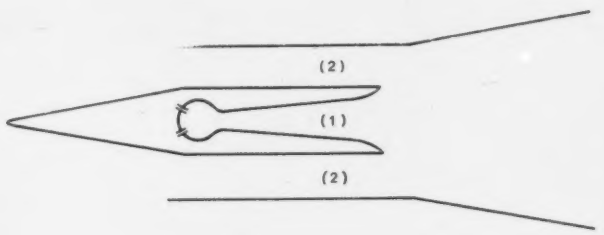
1. Electro-magnetic emission properties of solids.	Thermal emission of certain named reactor fuel and construction materials.
2. Aqueous solutions of organic liquids.	Aqueous solutions of a certain group of 60 named organic liquids.
3. Metal forming.	Extruding of titanium alloys for airframe construction.

Going back to the requestor can be time consuming and expensive and can easily be avoided by the user giving a precise statement of interest in the first place. A very clear and helpful actual request is shown in Fig. 10.

Purpose

If the purpose or reason for asking the question can be given then this affords just that little bit more background information which can help yield a greater insight into the real requirements. This request for purpose should not be confused with « need to know ». SDS does not operate on a need to know basis. The information, entirely up to the user to give or withold, is simply to put the request in its proper perspective.

Question/Purpose



Air-Augmented rocket propulsion system

Stream (1) is a rocket exhaust which still contains some unburnt fuel.

Stream (2) is air entrained from the atmosphere.

The system is axisymmetric.

Turbulent mixing and chemical reaction take place between the two streams with the release of heat.

Because the pressures are not equalized between streams 1 and 2 there will be shockwaves in the flow.

The problem is to find:

1) The location of any shockwaves in the flow.
2) The effect on the mixing of these shockwaves.
3) The effect on the mixing of simultaneous combustion.
4) In particular the effect of shockwaves and combustion on the turbulence properties, represented by the eddy viscosity.

References

NASA Tech. Note TND-4974. John P. Weidner and Carl A. Trexter. « Preliminary investigation of a momentum diffusion between two supersonic streams in the presence of shockwaves ».

AIAAJ Vol. 1, No. 11, Nov. 66, pp 2017 - Coleman du P Donaldson and K. Evan Gray. « Theoretical and experimental investigation of the compressible free mixing of two dissimilar gases ».

Descriptors

I	II
Shock waves	Jet flow
Shock wave interaction	Air jets
Normal shock waves	Two-dimensional jets
Oblique shock waves	Jet mixing flow
Combustion	Supersonic jet flow
Combustible flow	Turbulence
Supersonic combustion	Turbulence effects
	Turbulent diffusion
	Turbulent jets
	Turbulent mixing
	Turbulent wakes

Suggested Strategy

I and II

Fig. 10

References

Another useful aid in making sure that the most relevant material possible is obtained, is to give, if possible, one or two references to documents that have been of use in the scope of the enquiry. These will give a further indication of the subject and can provide additional search terms if they are between the time span of the system (1962 to date), since the files can also be searched by author or corporate source.

Keywords

The search will be conducted using keywords (or descriptors) which are logically combined. Any question defined by a paragraph written in clear language will be translated by SDS into the keywords used by the retrieval language. If, however, the user is able also to supply keywords relevant to the question, it will help very much for the query formulation. Two things should however be borne in mind; the first is still to write the question in the form of a short paragraph as mentioned above, the second is to take the descriptors if possible from the NASA thesaurus (NASA SP 7030) as this is the one used by SDS for indexing and searching.

Files

The database has been broadened to include not only the NASA file, but also that of the American Society for Metals (covering metallurgy and materials), Engineering Index (covering mechanical, electrical and civil engineering) and the US Government Clearinghouse (covering US R and D reports). Other files in different fields may be added to the database. The basic search will be carried out on the NASA file and the analyst will use his experience and knowledge to judge whether a search should be carried out on one or more of the remaining files.

Limitations

The following limitations are available in the system and may be useful in obtaining just what is desired:

- The system can limit a search by time, i.e. can give the result for any period in time from 1962 to date.
- It can give just report literature or only periodical articles as well as both.
- It can also give as the result a general survey with some irrelevant material or a specific search with little irrelevant material, or a very specific search risking losses.
- If the user can say approximately how many references he can cope with, the analyst will take this into consideration too, in which case only the most recent references will be printed.
- The references can be printed in two formats — full citation (normal) or just the accession number (eg 71N12345) which enables the abstract as well as citation to be found in the appropriate abstract journal (N = STAR, A = IAA).

These limitations are built into the SDS system. It is in the interest of the user to take advantage of them by stating just what he requires.

Request forms

As a guideline for question formulation SDS has prepared a request form (Q) in which the user is asked to give his name and address, to specify the kind of service he requires and to define the question, add further complementary information and give the search restrictions as they are described above.

To specify the type of service the user can choose between the following:

RB — Retrospective Bibliography.

SDI — Selective Dissemination of Information (12 monthly up-dates).

RB + SDI — Both required, i.e. an RB followed by an SDI.

SEARCH STRATEGIES

The co-ordination problem

The availability of a dialogue interrogation system to search interactively in the NASA file for document references relevant to a certain problem has greatly changed the time and effort involved in searching techniques. The philosophy behind the compilation of search profiles and the formulation of suitable strategies is still basically the same as with a conventional system, but it is the ability of the interactive system to do certain things that gives it its high placing.

The problem consists, for any question composed of more than one term, of finding the suitable logical combination of terms which will retrieve a satisfactorily high proportion of document references pertaining to the search question. It will be realised that any document is a combination of different concepts:

« Diffusion of carbon in high strength steels »
« Measurement of the tensile strength of titanium alloys »
« Evolution and motions of binary stars ».

In a co-ordinate indexing system these concepts would be translated into plain language keywords (descriptors) which could later be combined using Boolean logic to retrieve the original document. These keywords denote words selected from the title, abstract and text for indexing or retrieval purposes without consideration of the conceptual relationships to which they relate. Although pre-co-ordination and rôle-indicators are used, there is no means in the NASA Thesaurus language to distinguish between completely different concepts such as « Silver coatings on copper pipes » and « Copper coatings on silver pipes », both of which would be retrieved in a search on the simple terms copper, coatings, silver, pipes, in addition to various other possibilities. Such false-drops can, however, largely be eliminated by displaying the citations on the screen, thus ensuring that only relevant documents are retrieved.

Search logic

When conducting a search there are three logical operations from Boolean Algebra which may be employed and that indicate clearly the kind of combination required between the sets involved.

(a) *Logical product or Intersection*

This requires the simultaneous appearance in the index record of the descriptors concerned for the citation to be retrieved. This is the intersection proper and is represented by A **and** B (symbolically A * B).

(b) *Logical sum or Cumulation*

This requires the presence of at least one of the descriptors concerned in the index record for the citation to be retrieved. It is represented by A **or** B (symbolically A + B).

(c) *Logical complement or Exclusion*

This powerful device requires the absence from the index record of the descriptors indicated. The exclusion is represented as A, **but not** B (symbolically A - B).

A search strategy may include any or all of these logical operations, e.g. A + (B * C) - D. An important point to note is that the brackets must be included in an equation otherwise it will be wrongly interpreted by the computer. Consider A * (B + C): this clearly shows that Term A is to be intersected with **either** Term B, **or** Term C. If the brackets are omitted, the equation reads A * B + C and will be interpreted as Term A intersected with Term B and the result of this added to Term C, i.e. (A * B) + C.

On the other hand the equations should be kept fairly simple, i.e. at one level only; and with a maximum of 20 characters.

Thus although [(A * B) + C] * [(D + E) * F] would be accepted by the computer it is probably simpler to break it down. To leave it in the form (A * B) + C * (D + E) * F would yield an incorrect answer as outlined above. A correct formulation would be:

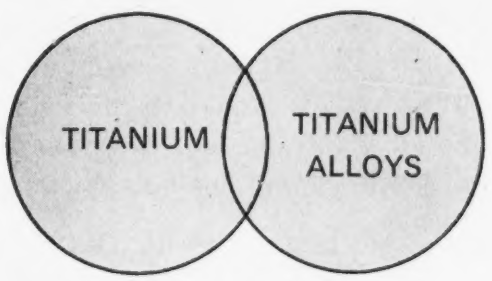
first to combine	(A*B) + C	= Set 1
then	(D + E) * F	= Set 2
finally	Set 1 * Set 2	= Set 3

The process of searching a subject file entails the manipulation of document classes identified by keywords (descriptors). The simplest search of all is that conducted on a single keyword, e.g. « Titanium », which will retrieve all document references indexed by that term. The class of acceptable documents can be enlarged by adding two or more sets, e.g. « Titanium » **or** « Titanium Alloys ». This indicates that the searcher is willing to accept any reference which has been indexed by either one of these terms or both.

Since searching on a single term is apt to retrieve irrelevant material, the class of acceptable documents can be reduced by an intersection, in which case the requirement is made that only citations assigned to the terms, say, « Titanium » **and** « Metal Forming » are to be retrieved. The set can be further reduced by using exclusion, i.e. negating a term. Thus it can be stipulated that documents dealing with methods of forming titanium **but excluding** extruding methods are required. (There may well be however some losses here if a document deals with both aspects).

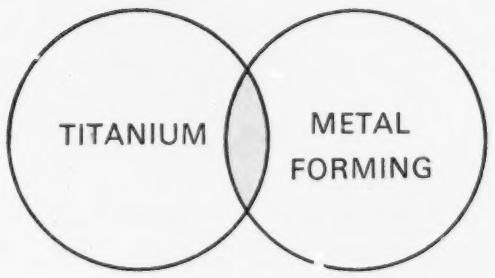
These formulations can be represented by means of Venn diagrams.

(a)



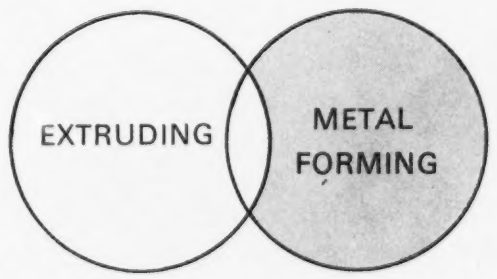
Titanium **or** (+) Titanium Alloys.

(b)



Titanium **and** (*) Metal Forming.

(c)



Metal Forming **but not** (—) Extruding.

Vocabularies

Having outlined the Boolean logic with which searches are conducted it is next necessary to consider the search (and indexing) language. It is unfortunate that in the SDS system one has to remember to employ two distinct vocabularies for searching the NASA file namely the NASA Subject Authority List (SAL) and the NASA Thesaurus (T).

The SAL vocabulary was in use from 1962-1967 and is a typical post co-ordinated thesaurus consisting of, in the main, simple concepts (or uniterms) expressed in the singular. Documents were indexed not only specifically, but also generically. Thus a document on « Spot welding of titanium alloys » would be indexed not only under « spot welding » and « titanium alloy », but also under the simple terms spot, welding, titanium, alloy. One could perhaps assume from this that by searching on the simple term « alloy » references to all types of alloy (e.g. titanium, aluminium, etc) could be retrieved. However, this is unfortunately not always the case and one must still search on the specific terms, e.g. « titanium alloy ». To be sure of getting the maximum number of references, it is advisable to search **also** on an intersection between simple terms, e.g. besides searching specifically on « satellite attitude control », more relevant citations can often be obtained by an intersection of « satellite **and** attitude **and** control ».

The NASA Thesaurus (T), in use since 1968, gives more pre-co-ordination of terms, as well as rôle indicators. It also expresses terms in the plural, thus alloys not alloy, resins not resin, and eschews the use of the adjectives employed in the SAL such as energetic, illustrative, high, low, small, etc. Documents are indexed specifically and the generic terms are not given in indexing. Thus a search on Metals or Alloys will **not** (as often in the SAL) pull out most of the different types of alloys and metals.

In the majority of cases, the corresponding term can be found in the SAL by dropping the « s » or some similar character from the T-term.

Thus:	**T Terms**	**SAL Terms**
	Steels	Steel
	X ray astronomy	X-ray astronomy
	Communication Satellites	Communications Satellite

Others can be found by searching the SAL for a similar word alphabetically

So:		
	Molecular Energy Levels	Molecular Energy
	Heat of Formation	Formation Heat
	Method of Characteristics	Characteristics Method

Some terms are completely different, e.g.

	Modulus of elasticity	Young Modulus
	Proving	Demonstration

but even these can be found with a little bit of effort.

To be sure, therefore, of being exhaustive in searching, care must be taken to include the SAL terms, as well as T terms, unless the search is to be limited to the years after 1967.

Generic relationships

Exhaustivity also relies on an expansion of generic terms. The Thesaurus gives the more specific terms of generic terms and also related terms. A search must take these into account as well as the SAL equivalents.

In any search concerned with the ionosphere in general, the concept of the ionosphere is represented by the following list of terms arranged hierarchically.

T Terms	SAL Terms
* Ionosphere	* Ionosphere
E Region	E Layer
E - 1 layer	E - 1 layer
E - 2 layer	E - 2 layer
E - 3 layer	E - 3 layer
* Sporadic E layer	* Sporadic E layer
* Lower Ionosphere	* Lower Ionosphere
D Region	D layer
* Upper Ionosphere	* Upper Ionosphere
F Region	F layer
F 1 Region	F - 1 layer
F 2 Region	F - 2 layer
	Lunar Ionosphere
	Night Ionosphere

* same term in SAL and T thesaurus.

If the expansion of narrower terms is effected by the « Expand Narrower Terms » feature of RECON, it should be noted that the lists contain, usually, only T terms (category 3) and terms used in both thesauri (category 2). In order to achieve sufficient coverage the T terms should be supplemented by their SAL equivalents (category 1).

In a general question such as « Flutter on Swept Wings », all the narrower terms for both concepts should be given to complete the coverage, e.g. subsonic flutter, supersonic flutter, swept back wings, delta wings, etc.

One does not, however, normally search on broader terms, that is to say, if a search on the specific subject « axial flow pumps » produced only a few hits, the searcher would be unlikely to search on the broader (generic) term « pumps », in the hopes that documents indexed by this latter term would be bound to deal with the former, since if a document **did** allow space to « axial flow pumps », then the precise terms should have been used in indexing.

Conceptual relationships

Broadening a search strategy is usually only done when the topic of interest is not adequately covered by specific terms in the Thesaurus, or when the subject falls into several hierarchies. Thus, in a search on « Aircraft stability in a two-dimensional boundary layer », the concept « two-dimensional boundary layer » cannot be expressed and the search should be broadened to cover the two aspects, i.e. « Aircraft stability in boundary layers » and « Aircraft stability in two-dimensional flow ».

The searcher will often be guided to related terms by the Thesaurus and these can also be used to broaden the search and pull in additional relevant material. Below are some typical examples :

- In a search on the « rolling movements of Spitfires », the term lateral stability should also be used.
- In a search on « heat transfer in circular cylinders », an additional legitimate search formulation would be « heat transfer in channel flow », since a cylinder is a channel.
- In a search on « Satellite boom deployment mechanisms », the searcher was guided to more search terms, not immediately apparent to him, such as « self-erecting devices » and « space erectable structures ».
- In a search on « fillers suitable for buoyancy applications », subsearches could be made on possible applications, e.g. life-boats, life rafts, etc.

A little ingenuity on the part of the searcher is required, especially in those cases where as in the last example above, relevant related terms are not given in the thesaurus, i.e. buoyancy does not give life-boats, etc. as related terms.

Concept grouping

One of the most important points in searching is to make an efficient grouping of concepts. Very often the division between concept groups becomes apparent only after a certain number of terms have been shown to be pertinent to the question. It is essential to group correctly in order to avoid a drastic reduction in recall. The following example will explain:

Consider a request for information on « Solar X-ray and ultraviolet line spectra ». This question consists of three concepts which have to be intersected to obtain the specific result:

1. the concept of origin - radiation originating from the sun
2. the concept of wavelength - X-ray and ultraviolet
3. the concept of scientific exploitation - spectral properties.

Below, some (not all) available thesaurus terms are listed in a matrix showing to which of the above concepts they refer.

	Solar Origin	Wavelength	Spectral Analysis
1. Solar X-rays	●	●	
2. Solar radiation	●		
3. Solar cosmic rays	●		
4. Solar Flares	●		
5. Solar Activity	●		
6. Chromosphere	●		
7. Oxygen spectrum			●
8. X rays		●	
9. X ray spectroscopy		●	●
10. Ultraviolet radiation		●	
11. Spectra			●
12. Spectral emission			●
13. Lyman Alpha radiation		●	
14. Ultraviolet Spectroscopy		●	●
15. Solar spectrum	●		●

These terms must be logically combined to obtain the desired result, and appropriate equations follow with comments:

(*a*) 1 * (11 + 12)

This gives solar x-rays **and** spectra **or** spectral emission. It is **not** necessary to include for example set 8 (x rays) or set 15 (solar spectrum) since these concepts are clearly covered already.

(*b*) (2 + 3 + 4 + 5 + 6) * (8 + 10 + 13) * (7 + 11 + 12)

Thus the solar terms are intersected with the wavelength and these two intersected with the spectra.

(*c*) (2 + 3 + 4 + 5 + 6) * (9 + 14)

Here the solar terms are intersected directly with concepts expressing both wavelength and spectral analysis and so there is no need to intersect again with the spectral terms. If one did, the result would be much smaller, since all three concepts would **have** to be present.

(*d*) 15 * (8 + 10 + 13)

Solar spectrum is here combined with the wavelengths to give the specific result; again then it

is superfluous and restrictive to put in other spectral terms as these have already been covered. The final result will be obtained by summing the four sets, i.e. a+b+c+d.

One further example will reiterate and perhaps clarify further this principle. In a search on « Rotational and Vibrational Spectra in Hydrocarbon Combustion », the following intersection would be made:

Hydrocarbon Combustion * Spectra (including narrower terms).

In addition the following intersection would also be made separately:

Hydrocarbons * (Flames + Combustion) * Spectra

The two results would be added to give the final one. Remember it would be pointless to put Hydrocarbon Combustion in the group with Hydrocarbons and to intersect with « Flames or Combustion » since the term already includes the latter concept. Narrower terms for each descriptor used would of course be added.

Omission of implicit categories

Repetition of categories can be avoided and also loss of references if concepts implicit in the question are not specified in the search strategy. For example, « Fluidics in Machine-Tool Control ». To the searcher, the question might appear at first glance to be described by three distrinct groups.

(Fluidics + NT) * (Machine-Tools + NT) * (Control + NT)

On closer investigation of references displayed on the screen, it transpires that wherever fluidic devices are used in conjunction with machine-tools, it will be almost certainly for control purposes and not to provide prime power. Thus no ambiguity, or very little, is created if the category of control is omitted. Consequently search time is reduced and the results more exhaustive.

Very similar is the dropping of a set because it is largely superflous, though not implicit. With regular use the searcher becomes aware of the number of citations of terms and which will result from a combination. Because of this, time can again be saved by making simple intersections and displaying output.

Thus, in a search on « miniature VHF antennas », where no Thesaurus term exists for VHF antennas, the normal formulation would be « miniature * VHF * antennas ». One would imagine that there would not be many references on this precise subject and so time and effort is saved by just combining the terms appropriate to « miniature » **and** « antennas ». The resulting citations can easily be displayed and those referring to VHF picked out and kept. Alternatively, the concepts « miniature » **and** « VHF » could be intersected and the antenna references pulled out. If the simple intersection **did** produce a large number of hits it is probably best and quicker to intersect with the third term rather than display all the citations. Note that it is the sets chosen to give the smallest number of hits that have been combined. It is probably obvious that there are likely to be far more references to « VHF antennas » than « miniature antennas » and so the combination « VHF * antennas » would not be done in the initial stages, if indeed at all. It is important to remember that this latter set « VHF * antennas » will include any references to « miniature VHF antennas », since this more specific concept must be included in the broader group. Similary « miniature * antennas » must also include all references to « miniature VHF antennas ».

To quote one other brief example, a search was requested on the « performance of airborne radar in a ground clutter environment ». To intersect all the parameters « performance * airborne equipment * radars * ground * clutter » would be bound to produce a negative response since each term must be present in the index record for the document to be retrieved. A useful result was obtained from a combination of « airborne equipment * radars * clutter », but more highly relevant references were retrieved by combining simply « airborne equipment * clutter ». In these references, the terms radar imagery and radar navigation had been used in indexing and although the titles indicated airborne radar performance, by using airborne radar on its own was obviously insufficient.

Term selection based on displaying

It can seldom be predicted which search strategy will be best to answer a question. As the concept of strategy includes the idea of feedback, the information supplied by the screen must be used and interpreted in order to choose a good strategy. Although searchers should have a fair idea of the strategy they will adopt when running the search, the formulation is bound to change in the light of the information obtained from the computer. It is of paramount importance that references retrieved in answer to a logical question should be scanned attentively in order to control the direction or scope of a question. To demonstrate this the following example is given ; a search was required on the « Pulfrich Effect » (this is an oculometric effect manifesting itself at different pupil dilations. As nerve propagation is a function of illuminance, a time difference arises in nerve signal transmission between the eyes, which is interpreted geometrically in the brain).

Several iterations had to be made to arrive at pertinent references :

1. (Visual acuity + NT) * (Pupils + eye examination) — no result.
2. (Visual acuity + NT) * (Perceptual speed + kinesthesia) — no result.
3. (Visual acuity + NT) * (Stereoscopy + NT) — no result.

It was impossible to look through all the terms on « visual acuity » to pick out the relevant hits as there were too many to make it feasible, but by displaying a few, possible terms could be selected and intersected to reduce the set. All combinations, however, yielded no relevant material, until the term « binocular vision » was intersected with « visual acuity ». As proof that these references were relevant, when they were displayed, they were found to contain the words « Pulfrich Effect » in the title.

Search modification

As mentioned before, practically all search strategies will be modified when it actually comes to running them on RECON. The search formulation can be varied in order to retrieve more or fewer document references as required. This can be done by moving up or down hierarchies, substituting synonyms and related terms, and by displaying the output. The logical statements of intersection (and), summation (or) and exclusion (and not) are the devices used to vary the exhaustivity of a search. It has to be remembered that by broadening a search, there will be the tendency to retrieve more irrelevant as well as additional relevant material.

Let us imagine the following request : « Supersonic Flutter on Delta Wings ». The NASA Thesaurus gives the following hiearchical breakdown :

Vibration
  Structural Vibration
    Flutter
      Supersonic Flutter

Airfoils
  Wings
    Swept Wings
      Sweptback Wings
        Delta Wings

By an intersection of the most specific aspects (supersonic flutter **and** delta wings), high precision will be achieved since the references will probably be highly relevant.

To improve recall, i.e. the search exhaustivity, by desiring a larger number of potentially relevant documents, the searcher can do one of two things : either he can reduce specificity or he can reduce exhaustivity.

By moving up the hierarchies to the more general classes, the search could be made on « Flutter of sweptback wings » or depending on which concept is the more important he could search on say « Supersonic flutter of wings » or « Structural vibration of delta wings ». Alternatively, the search could be made broader (less exhaustive) by searching on say the one term « Supersonic Flutter ». This would automatically retrieve any references on « Supersonic flutter of delta wings », but it would also retrieve a lot of irrelevant material (noise) such as « supersonic flutter on rotary wings » and so on.

Aspect searching

Often, a user is interested in various aspects of a question or such aspects appear during a search. With RECON, it is very easy to obtain these aspects and print them separately, at the same time ensuring that there is no repetition of the results.

As an example let us take the following request: « Desalination techniques, especially flash vaporization methods ».

Unfortunately, the term « Desalination » does not yet exist in the thesaurus, but « Flashing (Vaporizing) » does, so this can be selected, and it is found that there are only a few citations indexed with the term. Two other terms falling precisely within the scope of the question also appear with low postings. These are « Demineralizing » and « Demineralization », and can be used as independent terms like « Flashing (Vaporizing) ».

In direct answer to the question, an equation can be made between (Sea water, Brines + NT) on one side and (Evaporation, Purification, Treatment, etc.) on the other. On displaying the references obtained, other terms appeared to be relevant, and these can be intersected with the Sea water group to give an additional set of possible relevant references.

Our Set History is roughly as given below:

Set

1 - Flashing (Vaporizing)
2 - Demineralizing
3 - Demineralization
4 - Sea Water, etc.
5 - Purification, etc.
6 - Other terms (like Evaporation, etc)
7 - 4 * 5
8 - 6 * 4

9 - 1+2+3+7+8 OR PRINT 1

PRINT 9

9 2+3
10 9—1
PRINT 10
11 7—(1+10)
PRINT 11
12 8 —(1+10+11)
PRINT 12

Of course we could get the final result by adding all the relevant sets, but this would give a single set of references (set 9) containing both highly relevant and not so highly relevant references. Since the requestor is primarily interested in flash vaporization methods, we could print set 1 separately and all the other together, i.e. Sets 2+3+7+8. To keep the result in aspects however is done quite simply by negating sets already printed to avoid duplication.

Thus Set 1 is printed, then Sets 2 and 3 are added and Set 1 is subtracted from this result which is then printed. Similarly, the previously printed sets (1 and 10) are subtracted from Set 7 and Sets 1, 10 and 11 from Set 8 to give a result in four parts.

The parts are: Set 1 Flash Vaporization Methods used in Desalination
Set 10 Demineralization Methods used in Desalination
Set 11 Purification of seawater, i.e. desalination
Set 12 Other techniques of possible interest.

Each part thus deals with a different aspect of the problem and sets out separately (thus making for convenience) the relevant references.

Different approaches

No two searchers will have the same approach to a question, even though they will probably end up with a similar result. To give an idea of the two major types of approach — skeleton and exhaustive — a description will be given of two recent searches.

(a) *Skeleton searching*

This consists of doing an initial search and using the displayed results to give additional search terms.

The problem was « laser machining », and the strategy was obviously a simple intersection between (lasers + NT) **and** (machining + NT). The initial or skeleton search consisted of the following terms:

Lasers	Machining
Gas Lasers	Cutting
Semiconductor Lasers	Drilling
Solid State Lasers	Grinding (Materials Removal)
	Metal Cutting
	Metal Working
	Metal-Metal Bonding

After intersecting these groups, some 20 references were found. On scanning these, additional (SAL and T) terms were found such as

Pulsed Laser	Welding
Pulsed Lasers	Piercing
Laser	Melting
Laser Output	Spot Welding
Ruby Lasers	Metal Bonding

As can be seen the first intersection contained no mention of welding applications, which are of particular importance in laser machining. The results increased after the second intersection to 173 references.

After scanning some of the new references, still more additional terms were selected. This search history is not exhaustive, more results may have been obtained by using other types of lasers or coherent radiation. But as these terms did not appear while scanning references, it could be argued whether their use would yield more hits.

(b) *Exhaustive searching*

Another brief illustration will show a different approach to searching. The former example relied on picking up additional citations from the screen, but here an attempt was made to list **all** the relevant terms for use on RECON and then do the combinations. The subject was « Metal Forming » and some time was spent searching through the SAL and Thesaurus as displayed on the screen to pick out relevant terms. Thus terms concerning « Metal Working, Metal Forming, Metal Drawing, etc. » were entered as independent terms (i.e.they would be added (1 +2 +3,etc.) to the final results). Next all the various forming techniques such as rolling, cold drawing, extrusion, hot forming, stretch forming were found by paging the two vocabularies helped by the screen. These were summed to form one side of the logical equation. A similar method gathered the more common metals and alloys together with terms describing metal solids, e.g. sheet metal, metal strips, metal plates and so on. The intersection of these two major groups (forming **and** metals) was made and the result added to the independent terms to give a comprehensive bibliography on the required subject.

As previously noted, the result could have been given separately, i.e. the independent terms printed and then the intersected terms printed. To ensure no reference overlapped, the former set should be subtracted from the latter before printing the latter.

24

e.g. 1 Metal Forming + other independent terms

2 Forming Techniques

3 Metals and Alloys

4 2 * 3

5 4 + 1

PRINT 5

1

2

3

4 2 * 3

PRINT 1

5 4 — 1

PRINT 5

Of course, to use all possible relevant terms in the first instance can be not only time consuming, but also a waste, since many terms might not have brought additional hits, as the previous example showed (only a careful examination of the retrieved references would reveal this). It is up to the searcher to decide whether to aim for exhaustivity and comprehensiveness or to risk losing pertinent material in order to save time (and, course, costs!).

Note that in both approaches, maximum use was made of the display screen. In the first instance, using it to obtain relevant indexing terms from citations, and in the second, using it to obtain relevant indexing terms from the thesaurusal listing.

REMINDERS

1. If sets are added, then duplicate references, i.e. references appearing in two or more sets, are eliminated.
2. If a term is negated, then you risk losing pertinent documents which may have been indexed with that term.
3. Too many intersections will produce a very low result, since each set (or term) must be present in the index record. Thus you may lose relevant material because a document may not have been indexed with a term used in the intersection.
4. Put yourself in the indexer's shoes (and vice-versa) and think what he may have done — you'll be surprised!

RECON HARDWARE AND SOFTWARE

The RECON network

The RECON system, as operated by SDS, is a network in which analysts use remote consoles to communicate directly with a computer in order to carry out the search strategies.

To give a better understanding of the way in which the RECON system works, a description will be given of the programme in operation. A brief explanation of the equipment used in the RECON installation is first necessary.

Terminal equipment

Terminals are used to converse with the computer and consist of a cathode ray tube display with an alphanumeric keyboard. A record of the dialogue can be obtained from a teletype coupled with the console. Up to 16 modular remote consoles can be supported by one modem.

The control unit acts as a buffer memory which is required to transfer data between the computer and the terminal. When transmitting from the computer to the terminal it stores the information sent out at high speed from the computer to maintain the image display on the screen. For this purpose, it is equipped with a memory storing 1024 characters, 800 of which can be displayed on the screen (40 characters times 20 lines). The remaining character storage space is used for control commands. When transmitting from the terminal to the computer, the message typed on the keyboard is accumulated for rapid transmission after depressing the TRANSMIT key.

Data transmission cannot be effected by sending the binary coded information as d.c. pulses over the line. The bit patterns are used to modulate a carrier frequency in the appropriate rhythm, and this modulation is effected in a MODEM (MOdulator/DEModulator). On leased lines, transmission is effected in full duplex mode, that is, a separate two-wire line is used for transmission in each direction. In this case the carrier frequency is on continually in either direction. If only one two-wire line is used, such as in dialled connections, half-duplex transmission must be used. This means that between transmissions in either direction the carrier frequencies must be switched off and on, respectively. Transmission can be effected at different speeds; however 2400 baud (bits/sec) are preferably used although transmission at a lower speed (1200 baud) is sometimes necessary.

Computer equipment

The computer, an IBM 360/65, handles the system programmes and other system equipment. The amount of core storage required for an on-line information retrieval system will depend upon the efficiency of the system programmes and in RECON's case is 128k bytes when 4 or 5 terminals are in operation. The SDI programme requires only 90k bytes.

The computer handles programmes for jobs not connected with the RECON system at the same time. This capability is known as multiprogramming, and it increases the throughput of the computer. With multiprogramming, however, interruptions of jobs are random and a remote access user might be blocked while a job involving long calculations with little input/output monopolizes

the control processor. A solution to this problem and that used at ESOC is the interruption of jobs according to priorities. RECON is allotted the highest priority.

Another solution is the use of a time-sharing system. Here no one job is allowed to monopolize the processor and periodic interruptions are made to allow other jobs to proceed. Thus a remote user will have the impression of being the only user of the computer.

While the computer system used for RECON is not time-sharing, because of rapid response time, the remote user still appears to have sole use.

The multiplexer is an input/output channel at the computer which can handle up to 8 control units (each with other I/O units) in parallel. It can operate in two modes - bytes or burst.

Storage media.

The data files best suited for storing the information are random (or direct) access units such as magnetic discs, drums, core and cells rather than magnetic tapes. Tapes are a cheaper form of storage and easier to update, but they have to be searched sequentially, resulting in a slower access time. The RECON system makes use of magnetic discs and the data cell which is a very large capacity mass-memory composed of individual magnetic strips which can be separately extracted and read in order to provide direct access to a store of some 400 million bytes. The amount of « back-up » store provided by these units depends on the number of references on file and the rate of addition, which in NASA's case is 500,000 references and 6-7,000 additions per month.

The RECON system

An on-line system may be defined as one in which the input data or request is transmitted directly from a point which may be some distance from the computer and the output data is transmitted directly back to the same point without the intermediary operation of punching data onto cards or tape. The remote user can by means of the keyboard and display screen, submit a request for information to the computer, which will search its files for the relevant information, perform any necessary computations and transmit the results of the search back to the user via the display screen and teletype.

The way in which this operation works is described below:

- There are two main files used in the RECON system. These are the « Inverted File » (IF) and the « Linear File » (LF) and both are stored on the data cell.
- The Inverted File contains an alphabetical list of all the keywords (descriptors) in the NASA thesaurus giving for each one the NASA or IAA accession numbers of the documents in which that descriptor appears.
- The Linear File contains the full bibliographic description of every document on file (about 500,000) followed by every descriptor used to index it.
- The data cell contains in addition to the IF and LF a complete alphabetical list of related terms, i.e. narrower terms. These terms still, of course, appear in the IF.
- Rapid access to the descriptors and their accession numbers on the Inverted File is made possible by the « Index of Inverted File » (IX) which is maintained on magnetic disk. The index contains the descriptors in alphabetical order together with pointers to the Inverted File for each. Thus the pointer will direct to the descriptor followed by its accession numbers or the descriptor followed by its related terms. In addition the Index has also the number of accession numbers and the number of related terms for each descriptor.
- Since the new citations are just added on the end, the Linear File is not sorted into accession number order. Thus an index has to be maintained in accession number order with pointers to the correct location in the Linear File so that the document citation can be displayed on the console screen. This index is known as the « Index of Linear File » (LX) and is also stored on disk.

- The RECON master programme is contained on magnetic disk and when loaded into the computer generates a message saying « This is the ESRO RECON » and giving the date of the last monthly updating. This message signifies to the user that everything is ready and that searching can begin.
- Once the RECON programme is loaded, the computer will poll every terminal at regular intervals of about 1 second depending on how many terminals are operational. Any data that has been entered on the screen and transmitted to the computer by the depression of the TRANSMIT key will be taken by the computer on its poll. The word « Processing » signifies that the command is being executed by the computer.

A description of how to conduct a search and an explanation of results which appear on the screen are given elsewhere. The way in which the various files are accessed, though, may be of interest to some and this is given below.

- If a term, say SATELLITES, is expanded, then the Index of Inverted File on disk is first accessed to give an alphabetical display of terms and « SATELLITES » appearing as E6 (Fig. 7). With this display, the number of citations and also the number of related terms are given.
- If the related terms to SATELLITES are required to be displayed, then on expanding E6, the data cell is accessed (directed from the disk by a pointer) to give the list of related terms.
- To SELECT the term SATELLITES or E6, access is again made via the disk for speed and a set is generated for the term, together with the number of citations. At the same time, the data cell is accessed and the machine stores on disk the set together with the list of accession numbers relating to the term obtained from the data cell. The same thing occurs for any other selected term, thus when the COMBINE command is given the list of accession numbers can be matched according to the search logic.
- When a set is required for display (Fig. 9), the disk store is accessed which points to the Index of Linear File on disk which points in turn to the location of the citations on the data cell.
- If it is wished to display an accession number only, then the entry point is made to the Index to Linear File and pointed to the Linear File. Similarly, if the « Set History » is wanted, the sets which have been stored on disk will be accessed to give this.
- The author/corporate source are treated as if they were descriptors and can be used in exactly the same way as the sequence for « SATELLITES ».
- The END RECON command enables the desired citations to be printed by causing the list of pointers to the Linear File to activate the citations on the data cell.

Table 1 - RECON FILE AND INDEXES

Abbreviation	*Name*	*Function*	*Storage medium*
IF	Inverted file	Descriptors followed by accession numbers	Data cell
LF	Linear file	Bibliographic references followed by descriptors	Data cell
IX	Index to inverted file	Descriptors pointing to IF	Disk
LX	Index to linear file	Accession numbers pointing to LF	Disk
RT	Related terms	Principle term followed by its related terms	Data cell

Update Procedure

Magnetic tapes are sent each month from NASA to ESOC in the form of a linear and an inverted file. These tapes contain the latest document references and these must be added to the RECON master file. The procedure for doing this is as follows:

- The new Inverted File (IF) tape is merged with the master IF tapes.
- The new Linear File (LF) tape is inverted on authors and corporate sources and merged with the master author tapes and the master corporate source tapes respectively.
- All inverted files (IF, author, corporate source) are loaded on to a data cell. At the same time an Index of Inverted File (IX) is created on disk. The Related Terms are already on the data cell before the inverted file is loaded.
- The new LF is loaded on the data cell creating an Index of Linear File (LX).
- The new LX is merged with the old LX on disk.
- The inverted IF of update only is loaded on to the data cell and an IX created on disk.

SDI Procedure

After the update has been done it is then possible to run SDI's and other current awareness services. An SDI is first prepared using a terminal to achieve optimum search descriptors. These descriptors, together with identifiers are punched on to paper tape and then converted to card images on magnetic tape. There is one card image for each command, descriptor or identifier. The magnetic tape containing the profiles is fed into the computer and a search is conducted on the update file for each profile. The operation is in fact exactly similar to a normal console search except that here the batch mode is used. The retrieved references are stored on magnetic tape and then printed from this tape on to paper. Statistics are derived at the same time from the disk drive data.

DEVELOPMENT OF NEW FEATURES FOR RECON

During the time in which RECON has been operational in ESRO, several suggestions have been put forward which appear to have the possibility of increasing the potential of the system and of giving even better end-results. A brief description of the suggestions is given here. It is thought that these facilities, if developed, would be of great value in the searching of other data-bases which are not necessarily constructed around the present software.
It must be emphasised that these are only suggestions and that any trials being made are in order to test the feasibility of the proposals. Many additional features would be welcome, e.g. the ability to search by report number or the generation of author and subject indexes to bibliographies, but while these do not present great technical problems in themselves to achieve, the main drawback lies in the storage requirements.

Improvements to Boolean formulations

(a) *Published terms retrieval capability*

Although Boolean Logic is, in the main, efficient, false associations will often be retrieved when two or more terms are intersected. These can be eliminated to a great extent by displaying the references on the screen, but clearly this becomes uneconomical when a large number of references are retrieved in answer to a query.

There exists already a differentiation in importance between descriptors assigned to documents, namely the descriptor prefixed by an asterisk in the search printouts. The asterisked terms are known as « published terms » since they are the descriptors used in the subject indexes of the abstract journals (STAR and IAA) to describe a document succintly.

It has been suggested that by conducting a search on these published terms, noise would be eliminated to some extent and the remaining irrelevant references could be taken out by screening.

A trial is to be carried out to check the possibilities of this assumption.

(b) *Title-word retrieval capability*

Since concepts expressed by a document must be necessarily constrained by the index language, terms are then required which are not used in the thesaurus. Thus there is a very real possibility that a relevant document may be missed in a search.

Consider a recent search for documents on « convoluted tubing ». The term « convoluted » does not exist in the thesaurus and it was extremely difficult to find useful synonyms. It was equally impossible to scan through some 7,000 references on tubing and pipes to see if the word « convoluted » appeared in any title.

To help overcome this problem a title-word capability is being considered so that analysts may search on terms used in titles. It is thought that the use of floating stems will cut down quite considerably the amount of storage needed for such a facility.

These two capabilities — the ability to search on published terms and title words — require inverted files to be created which would then be merged with the existing alphabetic listing of terms. The resulting file would be equivalent to having a weighted indexing capability.

Free-language formulation

Consideration is also being given to the development of a free text or syntactical linking capability allowing the retrieval of natural language expressions in a defined field of a set of document citations.

Consider for example « impurities **of** iron **in** copper ». The normal Boolean equation would retrieve not only relevant documents but also irrelevant ones such as « impurities of copper in iron ». Searching on published terms would also retrieve irrelevant documents such as the latter example though possibly not as many as the pure logical statement.

It is thought that the irrelevant citations could be eliminated to a certain extent by the use of controlled free language terms such as prepositions or word associations and other morphemes. Thus pertinent documents in a retrieved set could be isolated by using an expression such as (in copper) taken from the title. This would reduce the number of citations to be screened to enable the remaining relevant references to be picked out, hence giving a faster response time.

Such a facility has already been developed in IBM's data processing system (DPS) and this might be adaptable for RECON. It is thought that the capability would be limited to a retrieval subset of not more than 500 citations.

Term extraction and frequency of occurrence

As another aid to retrieving pertinent documents a ranking procedure is being considered which will give the frequency of descriptors appearing in the index records of retrieved citations.

After an intersection, the terms in the resulting set of citations could be displayed by order of frequency, so that the analyst could see at a glance whether or not he has forgotten to include in his search strategy a high ranking term or whether a low ranking term could be safely excluded.

This procedure would obviously be useful when skeleton questions are used, since it might appear that there is no need to be completely exhaustive in the search formulation.

Document ranking

It has been suggested that since it can be shown that for each Boolean search there is an equivalent weighted formulation that gives a similar output, by using the extraction facility it is possible to reformulate a query semi-automatically to do this.

The screening (keep) operation on a sample from a retrieved set would be kept in two subsets — one positive (hits) and one negative (noise) (see below). The automatic term extraction from these two subsets would give the relative + or — weights of the most frequently used descriptors, and from this a cut-off threshold value could be established. These term weights would only be used for documents obtained by a broad formulation.

Keeping references

At the moment, when scanning, pertinent references can be kept in set 99. It has been felt for some time that a further useful feature would be the addition of another KEEP set which would be used to store those documents **not** relevant. This would enable the facility above to be developed. In addition however, when it is known that the majority of references in a set will be pertinent, then it is clearly very time-consuming during the scanning process, to select the relevant references for keeping in set 99. It would be much more useful to have another KEEP set where the analyst can save time by storing only the few non-relevant documents in order to negate them later from the final search result. This procedure obviously requires fewer commands to the computer.

File of related terms

The alphabetic listing of descriptors on RECON is a combination of two thesauri — the Subject Authority List (SAL) used by NASA from 1962 to 1967 and the NASA Thesaurus (T) in use since 1968.

In general the SAL employed terms in the singular, e.g. LASER and the NASA Thesaurus (T) uses terms in the plural, e.g. LASERS. This fact is indicated on RECON by the allocation of types to all the terms. Thus type 1 signifies an SAL term, type 3 is a T term, while type 2 means that the same is used in both the SAL and the Thesaurus.

At the present time, the related term (more strictly, narrower term) feature on RECON is arbitrarily limited to 94 terms in types 2 and 3 only. Thus if the term LASERS is expanded, a display is obtained of all the narrower terms — i.e. the types of lasers such as gas lasers, ruby lasers, and so on, all of which can be selected in **one** group, with only one command.

This hierarchy is not available for the type 1 (SAL) terms and so the analyst must not only select the term LASER but also gas laser, ruby laser, etc. all as **separate** items. It is therefore proposed to update the related term file to include not only new terms added since 1968 but also the narrower terms used in the SAL.

The concept of related terms will also be extended somewhat to include in certain cases, the true related terms where no narrower ones exist. For instance, under the term CYLINDERS an instruction is given to use a more specific term such as circular cylinders, rotating cylinders, etc. If this term (i.e. CYLINDERS) is expanded these types of cylinder are not displayed in their hierarchy. Consequently if all types of cylinders are required they must all be added singly. NOZZLES is another case in point. The updating of the RT file would ensure that in these cases the more specific terms are added. It is hoped eventually to have a permuted index which will again facilitate term selection.

Output presentation

Five record formats are currently available:

accession numbers only (format 1); full citation including descriptors (format 2); accession numbers, title and NOC only (format 4); the full citation but excluding the descriptors (format 5); and a special file dump (format 3). Consideration is also being given to another format - the full citation including abstract.

To have the abstract of a document would clearly be most helpful and save the user considerable time. Since it is impossible (for reasons of storage space) to store all the abstracts in the computer, (although perhaps the current year only could be considered) some system such as the « CARD-compact automatic retrieval display » might be utilised whereby abstracts stored on microfiche could be called up on the RECON display screen and the printed as necessary.

Format 5 is useful for bibliographies which have been particularly carefully prepared, where the index terms would add nothing to the potential relevance of the citation and where the user does not require them to be present. Such a case might be when a search has been conducted on the published terms (see above). Here the retrieved citations would be considered to be of high pertinence and thus printing time and paper space could be saved by omitting the index terms entirely. The current awareness services could also be printed in this format.

It is also possible to display up to four references at once in format 4 with only one display (%) command.

RECON COMMANDS

To be able to use RECON, the analyst must know the appropriate commands which are based on symbols [1] from an ordinary typewriter keyboard. Figure 4 is a picture of the CCI RECON keyboard. The top two rows are the function or command keys and these are described below. Execution of a desired command is accomplished by depressing the desired command key. This action will cause the special character representing the command to appear on the lowermost line but one of the display screen immediately following the computer-written word « ENTER ». After entry of a command symbol, the desired data (descriptors, set numbers, etc.) is keyed and will appear on the screen after the command symbol. Finally, the « TRANS » (transmit) key [2] is depressed to send the data to the computer. The transmit key must be pressed each time after entering a command and the relevant data. If a mistake is made the cursor, which is the tiny line appearing beneath the space where a character is to be typed, can be retraced by means of the arrowed keys to the error and the correct letters or figures can be immediately typed in. The cursor must then be taken back to its original position.

BEGIN SEARCH (!)

The BEGIN search command will produce the preliminary questions relating to a search (Fig. 6), i.e. title, analyst, requestor, address and file. One line is allowed for the title and two for the address. It should be noted that file 1 is the NASA retrospective file, i.e. 1962 to date, and file 2 is the NASA update file. File 3 is the Metadex file, file 4 is the Engineering Index file and file 5 is the Databank on Electronic Components test file. Other files will be introduced later such as USGRDR, CAC etc.

BEGIN SEARCH BYPASS (!B)

If used, the preliminary questions will not be produced and the search will be conducted on File 1. If any other file is required, the normal BEGIN search command must be used, and where it says « ENTER NUMBER OF DESIRED FILE » then the appropriate number should be entered.

FILE SWITCHING

It is possible in the course of a search to switch from one file to another by entering the command . FILE 3 (where 3 is the desired file - in this case Metadex). A collection reset is performed, but the set numbers continue to run consecutively. It should be remembered that only 98 sets are permissible (set 99 being the keep set).

Although it is possible to combine and print sets taken from two or more different files, this is not advisable at present due to the accession number irregularities. It is preferable to give a « PRINT » command followed by « END BYPASS » **BEFORE** switching files. This will ensure that the references from each file are printed out correctly.

---

1. *The command symbols given throughout are those used on CCI equipment. The IBM commands (see appendix) are the same, but the symbols are different. It should be realised that the actual commands below are those* **within** *the parentheses and do not include the brackets themselves i.e. the Begin Search command is simply! and not (!).*

2. *It should be noted that on some terminals the INT (interrupt) key is used instead of the TRANS key. In either case it is the second key in from the left that must be depressed.*

Fig. 11

EXPAND (")

When this command is entered and followed by a word or part of a word alphabetically close descriptors prefixed by the letter E and a running number will be displayed on the screen relating to the word expanded, which will be numbered E 06 (Fig. 7). If a term displayed has related terms, then these too can be seen by expanding the E number of the term, e.g. "E6 (the 0 need not be inserted), (Fig. 11). The descriptors in this second display will be prefixed by the letter R and a running number, and are narrower terms to the keyword expanded. The term expanded will be R1 (not R6).

SELECT (≠)

When this command is entered followed by a word, the term selected will be printed out with a set number and the number of citations in which it appears. The E and R numbers may also be selected to avoid typing long words, e.g. E6. It is not necessary to type the 0 which appears in E01 to E09 and R01-R09. If all the terms are wanted, e.g. all the narrower terms, then they can be selected as: R1-R10, in which case just one set will be generated. In addition an indication of the coverage is also given after the set

e.g. 3 125 E6-E12
E6: Space Flight

and will always be either E6 or R1, depending on which group is selected. Groups as well as individual numbers can be selected by using a comma, or a hyphen, e.g. E8, E11, E13-E16, E20, R1-R4. The letter E or R must always be inserted. If when a word is selected the computer says « Term not in dictionary » by pressing just the EXPAND key and then the TRANSMIT, key, the rejected word will be displayed on the screen in alphabetically ordered display. It should then be possible to select the correct word. This is still possible to effect even if the word has been selected correctly.

STEM SEARCHING (?)

It is possible to select groups of terms with the same root or stem using the truncation sign (?). If for example the command ≠ DIFFUS? was given then all terms with this stem would be selected and combined in one set i.e. all terms from E2 to E19 in figure 11 would be selected. Attention must be paid to the use of this procedure since there is a risk not only of easily reaching a disk storage overflow but also of selecting non-relevant material.

COMBINE ($)

Sets may be combined (see page 19) by using the above command. To avoid typing a lot of numbers, consecutive sets can be compressed, e.g. $ 1-9/+. To avoid truncation of the record of sets chosen due to screen line-length limitations, it is best to combine no more than 10 sets. In some combinations, e.g. where 19 sets have already been selected, set numbers, time, and more important, storage, can be saved by adopting the following procedure:

Combine sets 1 to 19. Although this can probably be done (depending upon the number of references in each set), it is undesirable for reasons such as shown in the paragraph above. Thus it is best to limit this to combine sets 1 to 10 to generate set 20. Then, instead of combining sets 11 to 19 to give set 21 and adding 20 + 21 to get 22, it can be done in one step by combining sets 11 to 20, and generating set 21.

INTERSECT

In order to see at a glance whether a logical combination will be worthwhile the command . INTERSECT can be used as it enables the comparison of two groups of sets. It also can cut down on the amount of storage space required since it gives an indication of how many references will result in a given combination without actually storing them.

The command . INTERSECT (1, 2, 3, 4) * (5, 6, 7) will give the logical intersection of the sets of the first parenthesis with those of the second. The logical operation $ 1*5, $ 1*6, $ 1*7, $ 2*5, $ 2*6... will be performed systematically and the results presented in the form of a matrix which can easily be read.

DISPLAY (%)

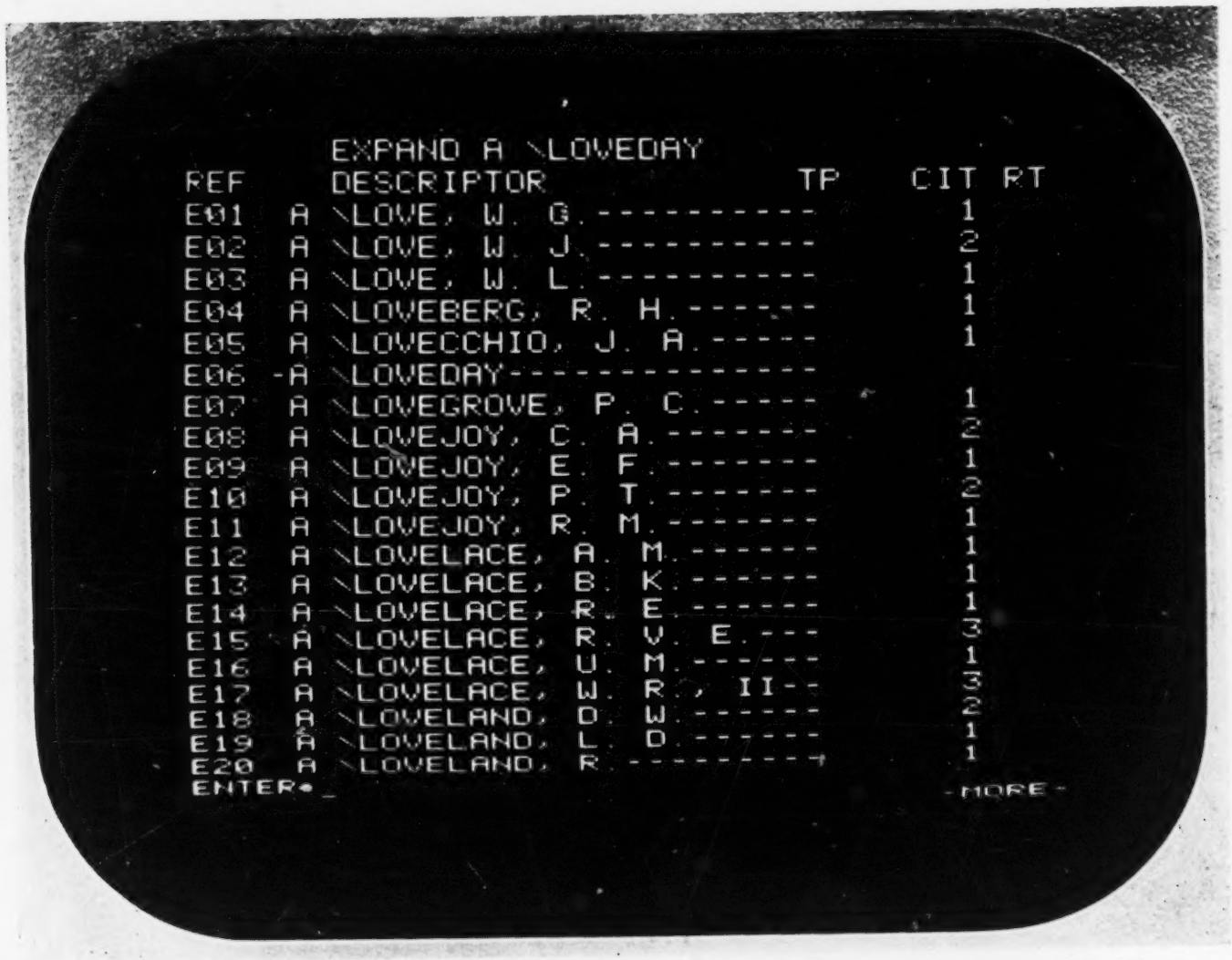
This command, followed by a set number, e.g. % 3, causes a display (Fig. 9) of the citations in that set, the most recent first. The next reference can be displayed by simply pressing the DISPLAY button again, followed by the TRANSMIT button (on the IBM 2260 the item command is used for displaying subsequent citations). There is no need to enter the set. By entering % — the preceding citation will be displayed again. A particular citation (say number 49 in Set 3) can be displayed without going through the preceding 48 by entering % 3/2/49. In this case, however, preceding or subsequent references cannot be displayed unless the command to display, say 50-60 in set 3 is given to break the sequence e. g.%/3/2/50-60. The second number stands for the format — where format 1 is accession number only and format 2 is the full citation. Other formats may be added later. The format number must be inserted when using this form. A specific reference may also be displayed by simply entering the accession number e.g. %70N12345.

PAGE (0)

This command is used when the word « MORE » appears at the lower right of the screen, i.e. in display, set history and expand displays. With regard to the latter each descriptor is prefixed with a running number labelled from E01 to E99. This code series is regenerated on subsequent pagings when E99 is reached. The process continues without limitations. As with the DISPLAY command, back paging can be done by using the command followed by the minus sign (0-).

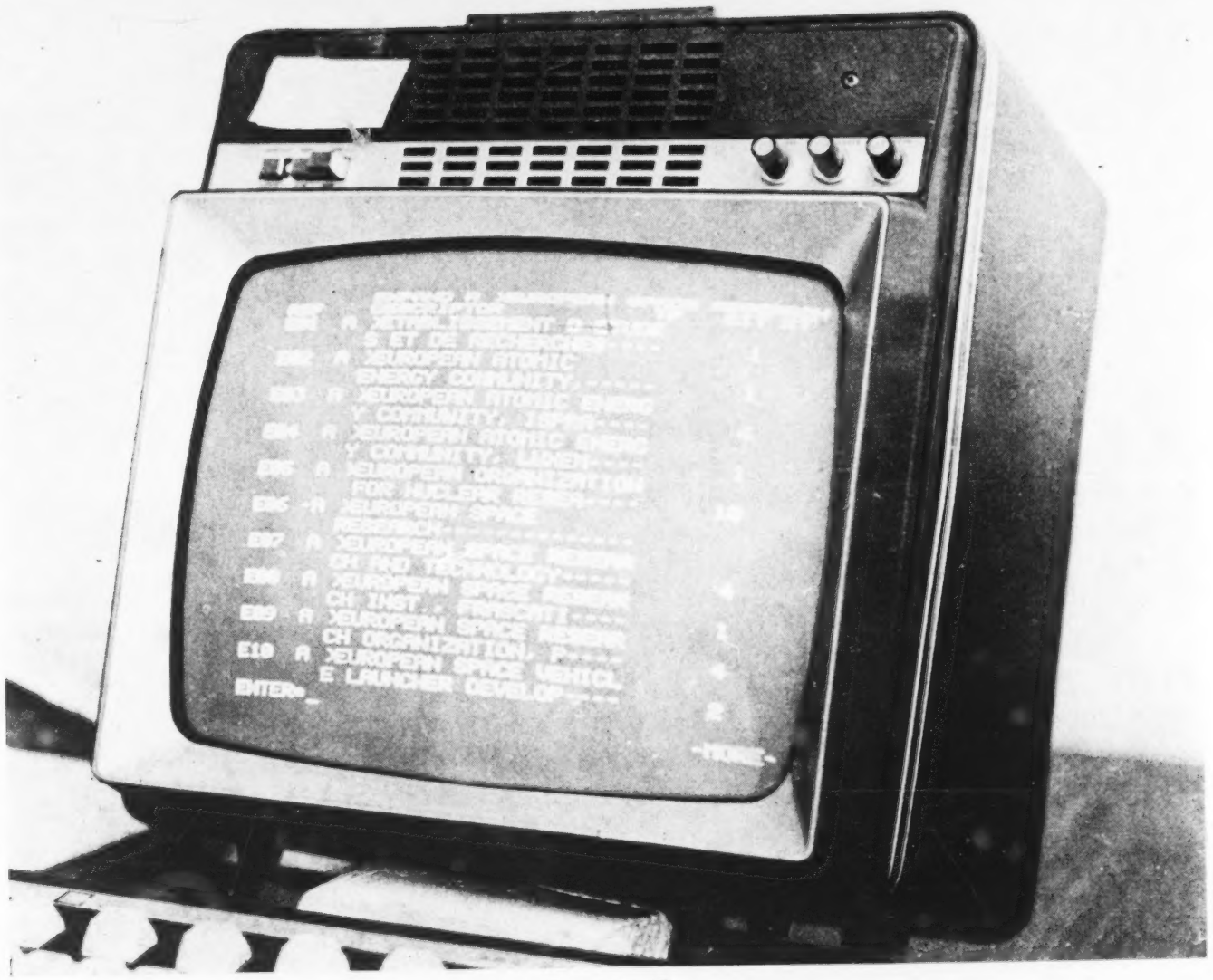
LIMIT ())

To limit a set, the above command should be used followed by the set, year, type and range documents required eg. 16/65-69/ALL/ALL (or simply)16/65-69). This means that set 16 is


EXPAND A \LOVEDAY
REF DESCRIPTOR TP CIT RT
E01 A \LOVE, W. G. 1
E02 A \LOVE, W. J. 2
E03 A \LOVE, W. L. 1
E04 A \LOVEBERG, R. H. 1
E05 A \LOVECCHIO, J. A. 1
E06 -A \LOVEDAY
E07 A \LOVEGROVE, P. C. 1
E08 A \LOVEJOY, C. A. 2
E09 A \LOVEJOY, E. F. 1
E10 A \LOVEJOY, P. T. 2
E11 A \LOVEJOY, R. M. 1
E12 A \LOVELACE, A. M. 1
E13 A \LOVELACE, B. K. 1
E14 A \LOVELACE, R. E. 1
E15 A \LOVELACE, R. V. E. 3
E16 A \LOVELACE, U. M. 1
E17 A \LOVELACE, W. R., II 3
E18 A \LOVELAND, D. W. 2
E19 A \LOVELAND, L. D. 1
E20 A \LOVELAND, R 1
ENTER•
-MORE-

▲ *Fig 12.*

Fig 13. ▼



limited to the years 1965-1969 but there is no limit to the type of documents included. The type of document is 1 for A (IAA) references and 2 for N (STAR) references. Thus)16/65-69/2/ALL means that the user requires only N numbered documents for the period 1965-1969. Range is the span of accession numbers required to limit the type further, e.g.) 16/65/2/12400-12500. If documents only for a given period are required then the command LIMIT ALL/65-69/ALL/ALL given at the start of a search before the terms are actually selected will limit every set in the search to the years 1965-1969. The command can be relaxed by entering LIMIT ALL/ALL/ALL/ALL.

KEEP (()

Citations may be kept in a set arbitrarily numbered 99, (this means a search cannot be taken beyond 98 sets). There are several ways of keeping references, but all use the command above. If it is necessary to keep a set, then the command is followed by the set number, e.g. (6. If references are actually being viewed on the screen then the one on display may be kept simply by depressing the KEEP and TRANSMIT keys.

The citation could also be kept by using the set number, format, and reference number, e.g. (49/2/7 which means that reference 7 of set 49 will be kept in set 99 in format 2. Similarly a group of references can be kept e g. (49/2/12-16.

SET HISTORY @

In order to cause a display of the search record, (Fig. 8), this command is used. The display will be the same as the teletype print-out, except that it will not show errors such as « Term not in dictionary », « Invalid Command », etc., and sets which when combined have produced nil hits.

AUTHOR DISPLAY ("A /NAME)

To display an author, (Fig. 12), the command and operation is the same for EXPAND, except that the EXPAND command here is followed by the letter A then a space, a slash and the surname, e.g. "A /LINDSAY. An alphabetical listing of all authors is then obtained with the chosen author as E6. Only one initial is given. This can be selected by the E number prefix to generate a set which can be combined in the normal manner with descriptors or other sets and which can also be displayed.

CORPORATE SOURCE ("A)NAME)

Similar to the above (Fig. 13), except that the right curved bracket is used instead of a slash, e.g. "A)ROYAL AIRCRAFT ESTABLISHMENT.

PRINT COMMAND (&)

Entry of this command followed by a set number will cause the contents of that set to be printed on a high speed printer at the end of the RECON session. The command has to be given again after each block of 50 citations. The format will be format 2. If format 1 (accession numbers) are required then the entry should be, e.g. & set/1/1-290. Accession numbers are printed in blocks of 500. Items being displayed can be printed by simply pressing the PRINT and then the TRANSMIT buttons.

TYPE COMMAND (')

To have a record locally, i.e. on the teletype, the TYPE command is used in the same way as the PRINT command. Format 2 citations are typed singly, while format 1 citations are typed in blocks of 84. An item on display will be typed out if the TYPE and TRANSMIT buttons are depressed.

END SEARCH (=) AND END SEARCH BYPASS (=B)

At the end of each search, the END SEARCH command if given will cause a display of questions (Fig. 14) to appear asking the user to express his satisfaction with the results and to make any comments he feels necessary. The answers to the questions and the comments appear on the search print-out. To avoid these questions the END SEARCH BYPASS should be used, which will then only give time elapsed for the search.

1-HIGHLY
2-MODERATELY
3-NO OPINION
4-DEFICIENT
PRESS INT AFTER EACH ENTRY
ARE RETRIEVED ITEMS RELEVANT=1
ARE THE RESULTS EXHAUSTIVE=2
ARE THE RESULTS WORTH THE TIME SPENT=1
WERE YOU INSTRUCTED IN USE OF RECON=1
IS THE MANUAL USEFUL=2
GENERAL COMMENTS (IF ANY)=MR RAITT PLEAS
E COMMENT ON USEFULNESS OF RESULTS_

Fig. 14

MULTIPLE COMMAND ENTRY

More than one command, even if completely different, can be entered at the same time by using the colon(;) as a separation sign and each will be executed consecutively. For example the following command « DIFFUSION »; ♯ E6; %1; %; % will be executed by expanding the term « DIFFUSION », selecting it and then displaying the first three references on the screen. This saves a lot of time especially when displaying references or switching files as &19; &; =B; .FILE 3.

It should be noted that it is not possible at present to stack « COMBINE » commands in this way.

MESSAGES (]No/)

To send a message to another terminal, the right hand square bracket is used followed by the number of the terminal, then a slash and then the text.

e.g.]3/Are you working today?

Terminal numbers are:-

1. SDS Paris (Master)
2. ESOC (Operators Console)
3. ESTEC
4. ESOC (Library)
5. SDS Paris
6. Rome
7. London
8. CNES
9. Stockholm.

If 0 is used as the terminal number, then the message will be transmitted to **all** terminals. There is a limit of two lines for the message.

SELECTED PRINTOUT FROM DISPLAY SCREEN

It is sometimes useful to be able just to have a certain portion of the displayed citation printed out on the teletype, instead of the whole. This can be done by making the END sign after the appropriate word by depressing the key END. The cursor must then be taken back to the beginning of the section before printing, out on the teletype. The cursor must then be reset to the correct position, i.e. after ENTER in order to continue searching.

APPENDIX - IBM TERMINAL

As already noted the command keys on the IBM terminal (Fig. 15) are different to those on the CCI terminal. Fig. 16 shows an IBM keyboard with the German version — the English keyboard is different only in so far as certain characters on the keys — the command positions are exactly the same.

Fig. 15

Fig. 16

```
        EXPAND BEARINGS
REF     DESCRIPTOR              TP    CIT RT
E01   BEAMS (SUPPORTS)--------3    456  5
E02   BEAMSHAPING------------1     19  1
E03   BEARING----------------2   1460
E04   BEARING (DIRECTION)-----3     72
E05   BEARING ALLOYS---------3     25
E06  -BEARINGS--------------3    179  7
E07   BEAT-----------------1     107  1
E08   BEAT FREQUENCIES-------3     98
E09   BEAUFORT SEA----------1      2
ENTER>,
```

```
        REF    DESCRIPTOR              TP    CIT RT
E10   BECKER-KISTIAKOWSKY-WILS
        ON EQUATION----------1             2
E11   BED-------------------1           267
E12   BED REST--------------2           169
E13   BEDDING EQUIPMENT-----3             1
E14   BEDIASITE-------------1             9
E15   BEDS------------------3             5
E16   BEDS (PROCESS
        ENGINEERING)---------3            56
E17   BEE-------------------1             1
                                       -MORE-
```

Fig. 17

```
          DISPLAY 02/2/1
70N21784A ISSUE 09 CATEGORY 15  TIL/DT/9089 LB/G-2761  00/07/68  UNCLASSIFIED RE
PORT
  THE FABRICATION OF TRITIATED TARGETS AT SACLAY  (FABRICATION OF TRITIUM TARGE
TS BY VOLATILIZATION AND ADSORPTION ON TITANIUM)
  MANIM, A.
  UNITED KINGDOM ATOMIC ENERGY AUTHORITY, HARWELL /ENGLAND/.  (UN878125)
  INFORMATION BRANCH.  DATE- JUL 1968  COLL- 10 P  TRAN- TRANSL. INTO ENGLISH F
ROM PROCEEDINGS OF THE EURATOM SYMP. ON ACCELERATOR TARGETS DESIGNED FOR THE PRO
DUCTION OF NEUTRONS /LIEGE/, 18-19 SEP 441967 AVAIL- CFSTI
  / ADSORPTION/*FABRICATION/ FILMS/*PARTICLE ACCELERATOR TARGETS/ RADIOACTIVE IS
ENTER>,                                                              -MORE-
```

Fig. 18

Before a command can be transmitted to the computer, a START symbol must be made. The command is then entered normally and the INPUT key depressed (with the shift key) to transmit the data. The START symbol and the command can be made anywhere on the screen — it does not need to immediately follow ENTER as on the CCI terminal.

Another interesting point is the word « PROCESSING » does not appear, instead the cursor moves from one end of the entered command to the other.

Otherwise, the procedures are the same as for the CCI equipment. Fig. 17 and Fig. 18 illustrate the difference in presentation between CCI and IBM for the EXPAND and DISPLAY commands.

CONCLUSION

A description has been given of the system used by SDS to disseminate combined and accumulated knowledge, as widely as possible, throughout Europe. The RECON network, with the full support of NASA, has gradually been extended so that centres in Member States, may, by installing their own terminals, have direct access to the data base. The data base itself has been broadened to include not only the NASA file, but also that of Metals Abstracts, Engineering Index and the US Government Research and Development Reports. Other files are under consideration. The feasibility of developing new features which would be of value in the searching of the data-base is being studied and a continuous monitoring of the system, services and procedures is taking place to ensure efficiency and effectiveness.

ESRO SP-63
European Space Research Organisation
SPACE DOCUMENTATION SERVICE :
OPERATIONS HANDBOOK
D. I. Raitt
45 pages
May 1971

I. Raitt, D. I.
II. ESRO SP-63

III. Text in English

An account is given of RECON, the on-line, interactive information retrieval system made available to ESRO by NASA. This system provides a means for describing a search interest to a computer, for obtaining displays and printout of results and for redefining the search based on examination of intermediate results. The search routine is described and types of search strategies discussed. The necessary hardware and software are briefly considered and proposals are made for development of new features to improve results.

ESRO SP-63
European Space Research Organisation
SPACE DOCUMENTATION SERVICE :
OPERATIONS HANDBOOK
D. I. Raitt
45 pages
May 1971

I. Raitt, D. I.
II. ESRO SP-63

III. Text in English

An account is given of RECON, the on-line, interactive information retrieval system made available to ESRO by NASA. This system provides a means for describing a search interest to a computer, for obtaining displays and printout of results and for redefining the search based on examination of intermediate results. The search routine is described and types of search strategies discussed. The necessary hardware and software are briefly considered and proposals are made for development of new features to improve results.

ESRO SP-63
European Space Research Organisation
SPACE DOCUMENTATION SERVICE :
OPERATIONS HANDBOOK
D. I. Raitt
45 pages
May 1971

I. Raitt, D. I.
II. ESRO SP-63

III. Text in English

An account is given of RECON, the on-line, interactive information retrieval system made available to ESRO by NASA. This system provides a means for describing a search interest to a computer, for obtaining displays and printout of results and for redefining the search based on examination of intermediate results. The search routine is described and types of search strategies discussed. The necessary hardware and software are briefly considered and proposals are made for development of new features to improve results.

ESRO SP-63
European Space Research Organisation
SPACE DOCUMENTATION SERVICE :
OPERATIONS HANDBOOK
D. I. Raitt
45 pages
May 1971

I. Raitt, D. I.
II. ESRO SP-63

III. Text in English

An account is given of RECON, the on-line, interactive information retrieval system made available to ESRO by NASA. This system provides a means for describing a search interest to a computer, for obtaining displays and printout of results and for redefining the search based on examination of intermediate results. The search routine is described and types of search strategies discussed. The necessary hardware and software are briefly considered and proposals are made for development of new features to improve results.